Received: 15 April 2020 | Revised: 13 August 2021 | Accepted: 6 September 2021

DOI: 10.1111/1460-6984.12673

International Journal of Language & Communication Disorders

RESEARCH REPORT

Statistical word learning in Catalan–Spanish and English-speaking children with and without developmental language disorder

Nadia Ahufinger[1,2,3] | **Laura Ferinu**[1,2] | **Mònica Sanz-Torrent**[2,3] | **Llorenç Andreu**[2] | **Julia L. Evans**[4]

[1] Estudis de Psicologia i Ciències de l'Educació, Universitat Oberta de Catalunya, Barcelona, Spain

[2] NeuroDevelop eHealth Lab, eHealth Center, Universitat Oberta de Catalunya, Barcelona, Spain

[3] Departament de Cognició, Desenvolupament i Psicologia de l'Educació, Universitat de Barcelona, Barcelona, Spain

[4] Department of Speech, Language, and Hearing, University of Texas at Dallas, Richardson, TX, USA

Correspondence

Nadia Ahufinger, Grup de Recerca en Cognició i Llenguatge (GRECIL), Estudis de Psicologia i Ciències de l'Educació, Universitat Oberta de Catalunya i Universitat de Barcelona, Barcelona, Spain.
Email: nadiahufinger@uoc.edu

Abstract

Background: A growing body of work shows that children with developmental language disorder (DLD) perform poorly on statistical word learning (SWL) tasks, consistent with the predictions of the Procedural Deficit Hypothesis that predicts that procedural memory is impaired in DLD. To date, however, SWL performance has not been compared across linguistically heterogeneous populations of children with DLD.

Aims: To compare SWL performance in a group of age, sex and non-verbal IQ-matched Catalan–Spanish and English-speaking children with and without DLD.

Methods & Procedures: Two cohorts of children: (1) 35 Catalan–Spanish-speaking children with DLD (M_{age} = 8;7 years) and 35 age/sex-matched typical developing (TD) children (M_{age} = 8;9 years), and (2) 24 English-speaking children with DLD (M_{age} = 9;1 years) and 19 age/sex matched TD controls (M_{age} = 8;9 years) completed the tone version of a SWL task from Evans et al. (2009). Children listened to a tone language in which transitional probabilities within tone words were higher than those between words.

Outcomes & Results: For both Catalan–Spanish and English cohorts, overall performance for the children with DLD was poorer than that of the TD controls regardless of the child's native language. Item analysis revealed that children with DLD had difficulty tracking statistical information and using transitional probability to discover tone word boundaries within the input. For both the Catalan–Spanish and English-speaking children, SWL accounted for a significant amount of unique variance in Receptive and Expressive vocabulary. Likelihood ratio analysis revealed that for both Catalan–Spanish and English cohorts, children having performance ≤ 45% on the SWL task had an extremely high degree of likelihood of having DLD. The analysis also revealed that for the Catalan–Spanish and English-speaking children, scores of ≥ 75% and ≥ 70%, respectively, were highly likelihood to be children with normal language abilities.

Conclusions & Implications: The findings add to a pattern suggesting that SWL is a mechanism that children rely on to acquire vocabulary. The results also suggest that SWL deficits, in particular when combined with other measures, may be a reliable diagnostic indicator for children with DLD regardless of the child's native language, and whether or not the child is bilingual or monolingual.

KEYWORDS
developmental language disorder, specific language impairment, statistical word learning, cross-linguistic

What this paper adds

What is already known on the subject

- Although there is some disagreement, a small but growing body of work suggests that deficits in procedural memory, as measured either by motor sequencing (Serial Reaction Time—SRT) or SWL tasks, may be part of the deficit profile of children with DLD. To date, studies have not examined SWL across linguistically heterogeneous populations of children with DLD to determine if it is a unique clinical marker of the disorder.

What this paper adds to existing knowledge

- The results show that children with DLD, regardless of their native language, or whether the child is bi- or monolingual, have difficulties on SWL tasks, and that these deficits are linked to severity of the language disorder. Taken together, these results indicate that procedural memory deficits may be a core feature of DLD. This suggests that statistical-learning tasks using tone stimuli can also advance our understanding of statistical-learning abilities in children with DLD more globally.

What are the potential or actual clinical implications of this work?

- The current study shows that statistical-learning tasks using tone stimuli can be used in conjunction with standardized assessment measures to differentiate children with DLD from children with typical language ability.

INTRODUCTION

Developmental language disorder (DLD) refers to a group of children with a neurodevelopmental disorder characterized by the inability to master spoken and written language comprehension and production in the absence of intellectual disability, hearing loss or other medical conditions or syndromes known to cause language disorders (Bishop, 2014). Although numbers vary slightly across countries, in the United States about 7% of English-speaking school-aged children have DLD (Tomblin et al., 1997). Although much of the research in DLD has focused on preschool and school-aged children, studies now show that DLD persists into adolescence and adulthood (Catts et al., 2008; Conti-Ramsden et al., 2013; Durkin & Conti-Ramsden, 2007; Tomblin et al., 1992).

Recently, there has been interest in characterizing procedural memory deficits in children with DLD following

Ullman's Procedural Deficit Hypothesis (PDH) that posits that procedural learning deficits are the underlying cause of language impairments in these children (Ullman & Pierpont, 2005; Ullman & Pullman, 2015; Ullman et al., 2020). In one of the first studies to examine procedural learning in children with DLD, Tomblin et al. (2007) employed a serial reaction time (SRT) task and found that adolescents with a documented history of DLD evidenced atypical procedural learning patterns not evident in the typical controls. Follow-up studies are consistent with the findings of Tomblin and colleagues (Evans et al., 2009; Haebig et al., 2017; Kemény & Lukács, 2010; Lammertink et al., 2020; Lum et al., 2010, 2012, 2014; Plante et al., 2002). For example, a meta-analysis of studies using SRT tasks, conducted by Lum and Conti-Ramsden (2013) and Lum et al. (2014), revealed a consistent pattern of poor sequential learning in children with DLD. Similarly, a meta-analysis by Obeid et al. (2016) that included a broader range of procedural learning tasks (i.e., SRT, artificial grammar learning, probabilistic classification, etc.) also found that children with DLD showed significant impairments in procedural learning as compared with controls, with task modality (visual versus auditory) not being a variable that moderated their observed effect sizes.

Key to the PDH account of DLD is impaired procedural memory, a type of non-declarative/implicit memory. Non-declarative memory is not a single construct, but a term used to characterize a type of learning that occurs on an ongoing basis over multiple trials or exemplars and where learning is manifested in the gradual changes in performance or behaviour across these tasks. Because knowledge acquired via non-declarative memory is not available to conscious recollection or verbal expression, it is also referred to as implicit memory (cf. Squire, 1992, 1994; Squire & Knowlton, 2000). Procedural memory is implicated in the acquisition and control of motor and cognitive sequential habits and skills, and the computation and use of rule-based procedures, such as the concatenation of the sequential information (Mishkin et al., 1984; Poldrack et al., 1999; Squire, 1994).

The PDH links impaired procedural memory to the morphosyntactic deficits in DLD and to lexical deficits in these children (Ullman et al., 2020). The ability to hold in memory the sequential order of phonemes and syllables within the speech stream to implicitly track and compute the probabilities of adjacent sound sequences is an aspect of word learning that relies heavily on procedural memory. Typically referred to as 'statistical word learning' (SWL), this ability to track and compute sequential statistics in the stream of speech—transitional probability—has been linked to word learning and vocabulary acquisition in typically developing children (Graf Estes et al., 2007; Saffran & Graf Estes, 2006). Ullman et al. (2020) have recently extended the PDH model and suggest that children rely on procedural memory to gradually and implicitly segment sequences of syllables or phonemes from the speech stream to acquire the phoneme sequences of words.

Only a small number of studies have examined the link between impaired procedural memory and lexical deficits in DLD. Evans et al. (2009) asked whether children with DLD (ages 6;5–14;4) were able to maintain the order of sequences of syllables in auditory memory to track the transitional probability of the sequence of these syllables and use this information to implicitly discover word boundaries within the continuous stream of speech and if this ability was related to children's vocabulary. After 21 min of exposure to a stream of speech syllables where the transitional probability across CV syllables was the only cue to word boundaries, they observed that, unlike the age-match typical controls, the performance of the children with DLD was no different from chance. Importantly, for the purpose of this present study, in a follow-up experiment, Evans et al. (2009) asked whether the DLD children's failure to track statistical information across speech sounds extended to non-linguistic auditory stimuli as well. A subset of the children from the first experiment completed a second experiment where they listened to a non-linguistic version of the SWL task consisting of a stream of pure tones. After 42 min of exposure to the tone language, the performance for the children with DLD was again significantly poorer than the normal language controls and no different from chance. Follow-up analysis revealed that the children's ability to use transitional probability to discover the word boundaries within the stream of speech was significantly correlated with both receptive and expressive vocabulary, suggesting that failure to track and compute the transitional probabilities within the stream of speech may be a factor that underlies the vocabulary deficits seen in DLD.

A small but growing body of work appears to support the idea that in addition to poor performance on SRT tasks, children with DLD also perform less accurately than typical peers on SWL (i.e., Haebig et al., 2017; Mayor-Dubois et al., 2012). For example, a meta-analysis by Lammertink et al. (2017) revealed robust auditory–verbal statistical-learning deficits in children with DLD. Taken together, these findings suggest that implicit learning deficits in DLD may extend beyond sensorimotor sequential learning to include implicit memory for cognitive sequences as well, in particular in the auditory domain.

Current study

Critical to the PDH account of DLD is whether the findings from Evans et al. (2009) extend to a more linguistically heterogeneous sample, and if children's SWL ability differentiates individuals with DLD from children with normal language. The goal of the current study was to compare the SWL performance of a group Catalan–Spanish children with and without DLD with that of an English-speaking cohort matched on age, sex and non-verbal IQ. A second goal of the study was to determine the likelihood ratios of SWL to differentiate children with DLD from children with typical language abilities and to evaluate the potential clinical usefulness of the task as a culturally unbiased marker of DLD. If, as current research suggests, impaired auditory statistical learning is a characteristic of children with DLD, then the pattern of SWL learning observed for Catalan–Spanish-speaking children with DLD should mirror that of the English-speaking children with DLD. Similarly, if SWL is linked to aspects of the lexicon, then the relationship between SWL and receptive and expressive vocabulary should be the same for Catalan–Spanish and English-speaking children. Finally, if SWL performance is a sensitive measure of procedural memory deficits in children with DLD, then the likelihood ratio for a true positive rate (i.e., accurately ruling in/out DLD) also should be the same for the children with DLD regardless of their native language.

EXPERIMENT 1

Methods

Participants

A total of 70 Catalan–Spanish (CS)-speaking children (22 girls and 48 boys), 35 children with developmental language disorder (DLD-CS, M_{Age} = 8;9) and 35 chronological age- and sex-paired-wise-matched typically development children with normal language (TD-CS, M_{Age} = 8;9) participated in Experiment 1. The children with DLD-CS were recruited from institutions, organizations and schools around Catalonia. Children with DLD-CS were identified with the help of the Catalan Center of Resources for Hearing-Impaired People (CREDA), members of the Catalan service for school counselling and guidance (EAP), and Catalan Association of Specific Language Impairment (ATELCA), which work in conjunction with public and private schools throughout Catalonia to identify children with DLD-CS or children with language difficulties. The children with TD-CS were recruited from public schools within the larger Barcelona metropolitan area. All participating families completed an informed consent form and a background information questionnaire. A final report containing the results of all the tests administered to the children was given to the family as a token of gratitude for their commitment and contribution to the study.

All participants met the following inclusion criteria: (1) normal non-verbal intellectual quotient (NVIQ) ≥ 75; (2) normal hearing at 500, 1000, 2000 and 4000 Hz at 20 dB based on the American National Standards Institute (1997); (3) normal or corrected-to-normal vision; (4) normal oral and speech motor abilities by a certified speech–language pathologist; and (5) were native bilingual Catalan–Spanish speakers.[1] Children were excluded if parents reported: (1) a neurodevelopmental disorder, (2) emotional or behavioural disturbances, (3) frank neurological signs or (4) seizure disorders or use of medication to control seizures.

The children in the DLD-CS group had a formal diagnosis of language impairment or were in process to be diagnosed and were receiving speech–language services at the time of the study. The TD-CS children were at grade level in school had no history, or diagnosis, of language-learning disability and had never received speech and language services. To confirm participant's language status, standardized testing was completed by two trained researchers at the time of the study and included the Nonverbal Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 2004), and Clinical Evaluation of Language Fundamentals—Fourth Edition, Spanish (CELF-4-Spanish; Wiig et al., 2006):[2] (1) Core Language score, (2) Expressive Language score and/or (3) Receptive Language score. For the children with DLD-CS, either Core, Receptive or Expressive CELF composite scores were ≥ 1.5 SD below age level expectations. For the children in the TD-CS group, CELF composite scores were all at or above age level expectations (Table 1). Non-verbal IQ was within normal limits for all the participants and the two groups did not differ in age or in the number of females/males per group (22 females, 24 males). In addition, to investigate whether children's SWL is related to their receptive and expressive vocabulary, all children also completed the Peabody Picture Vocabulary Test, Third Edition, Spanish version (PPVT-III; Dunn et al., 2006), and the expressive vocabulary portion of the Spanish version of the Kaufman Brief Intelligence test (K-BIT-Voc; Kaufman & Kaufman, 2004).

Stimuli

Because the goal of the present study was to extend the findings from Evans et al. (2009) to a group of Catalan–Spanish children, the stimuli for the present study was the tone language used by Evans et al. (2009) and tone language 1 from Saffran et al. (1999). The tone language consisted of 11 pure tones taken from the same octave (starting at middle C within the chromatic set A, B, C, C#, D, D#, E, F, F#, G, G#). Each pure tone was created using the SoundEdit 16 sine wave generator (Adobe, San Jose, CA, USA) and was 0.33 s in duration. The tones were combined to create a total of six, three-syllable tone words. The tone words did not conform to rules of standard melodic or musical composition. Some tones appeared in only one word whereas other tones occurred in more than one word. For example, D occurred in four of the tone words whereas G# only occurred in one of the tone words (Table 2).

TABLE 1 Age and standardized scores for language and cognitive assessment measures for Catalan–Spanish (CS)-speaking children with developmental language disorder (DLD-CS) and typically developing (TD-CS) children

	DLD-CS (n = 35)			TD-CS (n = 35)			Comparison	
Variable	**Mean**	**SD**	**Range**	**Mean**	**SD**	**Range**	$t(68)$	p**-value**
Age (months)	105.34	21.27	66–155	107.80	21.26	67–153	–0.48	0.63
K-BIT (IQ)[a]	99.08	11.69	82–119	103.51	9.76	88–125	–1.72	0.09
CELF-CLS[b]	72.57	10.89	45–89	108.74	6.09	95–125	17.14	< **0.01**
CELF-ELS[c]	73.22	8.77	52–87	108.45	8.11	89–128	–17.43	< **0.01**
CELF-RLS[d]	77.45	10.19	59–97	105.82	5.55	94–118	14.45	< **0.01**
Concepts & Directions[e]	5.80	2.30	1–10	11.66	1.71	8–15	–12.09	< **0.01**
K-BIT voc[f]	77.14	11.65	53–96	106.40	10.20	83–127	–11.17	< **0.01**
PPVT-III[g]	77.80	11.87	55–105	106.25	12.59	83–127	–9.72	< **0.01**

Note: [a]K-BIT IQ: Kaufman Brief Intelligence, Spanish version: Non-verbal intelligence score (Kaufman & Kaufman, 2004) Scaled scores (M = 100, SD = 15).
[b]CELF-4 CLS = Spanish Clinical Evaluation of Language Fundamentals, 4th Edition: Core Language score (Wiig et al., 2006). Scaled scores (M = 100, SD = 15).
[c]CELF-4 ELS = Spanish Clinical Evaluation of Language Fundamentals, 4th Edition: Expressive Language score (Wiig et al., 2006). Scaled scores (M = 100, SD = 15).
[d]CELF-4 RLS = Spanish Clinical Evaluation of Language Fundamentals, 4th Edition: Receptive Language score (Wiig et al., 2006). Scaled scores (M = 100, SD = 15).
[e]Clinical Evaluation of Language Fundamentals (Wiig et al., 2006): Oral Directions Receptive Subtest Score (M = 10, SD = 3).
[f]K-BIT vocabulary = Kaufman Brief Intelligence, Spanish version: Expressive vocabulary score (Kaufman & Kaufman, 2004) Scaled scores (M = 100, SD = 15).
[g]PPVT-III = Peabody Picture Vocabulary Test, Third Edition, Spanish version (Dunn, Dunn & Arribas, 2006) Scaled scores (M = 100, SD = 15).
Significance of the p-values is shown in bold.

TABLE 2 Transitional probability of tone words and non-words

Stimuli	**Internal transitional probability**
Tone words	
GG#A	1.0
CC#D	0.75
D#ED	0.65
FCF#	0.50
DFE	0.42
ADB	0.37
Tone non-words	
AC#E	0.0
F#G#E	0.0
GCD#	0.0
C#BA	0.0
C#FD	0.0
G#BA	0.0

The six tone words were combined in a random order with no silent junctures between the words to create a 21-min continuous stream of tones where an individual tone word never occurred twice in a row. The transitional probability within the tone words ranged from 0.37 to 1.00, whereas the transitional probability across the tone word boundaries ranges from 0.05 to 0.60. This overlap in the transitional probabilities within and across the word boundaries occurred three times in the 30 across-word tone pairs resulting in this across word probability of 0.60. This 0.60 probability occurred when the tone word GG#A was followed by the tone word DFE, as the cross-boundary sequence AD also occurred in the tone word ADB. In addition to the six tone words, six tone non-words were created (Table 2). These non-words were made up of the same tones from the language tone inventory, but because they never occurred in that order during the exposure the internal transitional probability of the non-words was 0.0.

The same two-alternative forced choice (2AFC) test as in Evans et al. (2009) was used to measure children's ability to use transitional probability to discover the boundaries of the tone words within the stream of tones. The six tone words and non-words listed in Table 2 were paired exhaustively to generate a 36-item test. Each test trial consisted of a tone word plus non-word pair. The tone word occurred as the first member of the pair for half of the test items and as the second member of the pair for the remaining test items. The order of test items with the tone word as the first item in the pair was counterbalanced across the trials and the order of the trials was randomized. The test items were recorded onto a digital recorder and presented in the same fixed random order to all the participants.

Procedure

Similar to Evans et al. (2009), the children listened to the continuous stream of tones while colouring pictures. To avoid potential ceiling effects, the Catalan–Spanish study was designed to mirror that of Saffran et al. (1999), and the children heard the exposure stimuli for a total of 21 min. Prior to the task, children were told:

You are going to have about 20 min to colour. While you are colouring, some weird computer music will be playing

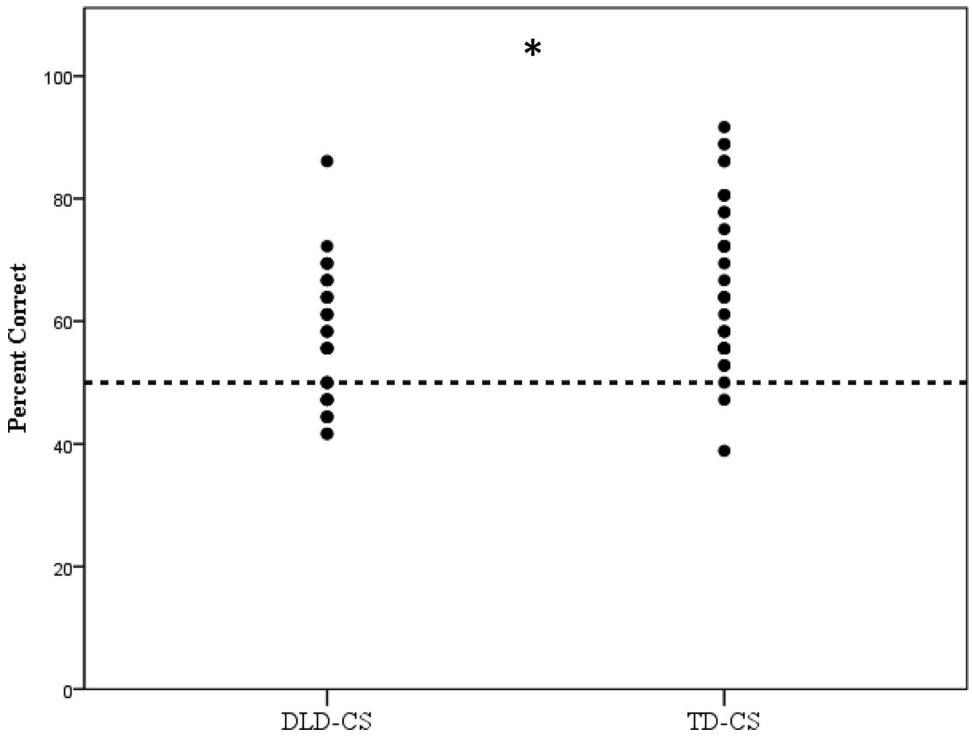

FIGURE 1 Percentage of correct answers for the tone version of the statistical word learning (SWL) task for the group of Catalan–Spanish-speaking children with developmental language disorder (DLD-CS) and the group of typically developing (TD-CS) children. *Note*: Chance equals 50%. *Significantly different between groups

in the background, but I would like you to focus on painting. When the music has finished, I will ask you some questions.

The examiner sat quietly behind the children during the task to ensure they were focused on the colouring task and were not distracted. At the end of the 21 min of exposure, children completed the test trials. Children were instructed to choose the sound sequence that sounded most familiar to the word tone stream from among two alternatives.

Prior to the testing phase, children completed a series of practice trials. Practice trials consisted of pairs of short melodies created from familiar Catalan–Spanish children's songs in a correct and incorrect order (e.g., the tune, without words, from 'Quan les oques van al camp' versus 'les van camp al quan oques'). For the practice trials children were told:

Now you are going to hear two sets of sounds and I want you to choose the set of sounds that sounds most like the weird computer music. If you do not know, it is ok to guess. First, we will practice. We're going to hear two different sets of sounds. I want you to tell me verbally if the 'one' set or the 'two' set sounds more like the song you know.

After completing the practice phase children completed the test trials. All the children completed all the practice trials successfully and no children were excluded from the study because they were unable to understand the task.

Results

Statistical word learning (SWL)

Accuracy for the DLD-CS and TD-CS groups are shown in Figure 1. The mean for the children with DLD-CS was 56.9% (SD = 10.3) and for the TD-CS group was 66.9% (SD = 13.5). Single-sample *t*-tests (two-tailed), where chance was 50% indicated that both groups' performance was significantly better than chance (DLD: $t(34) = 3.94$, $p < 0.001$; TD: $t(34) = 7.40$, $p < 0.001$). To determine whether the performance for the DLD-CS group differed from that of the TD-CS controls an analysis of covariance (ANCOVA) was conducted with non-verbal IQ as a covariate, because NVIQ approached significance between the two groups ($p = 0.09$). The result of the ANCOVA revealed that the performance for the children with DLD-CS was significantly poorer than that of the TD controls $F(1, 69) = 5.99$, $p < 0.01$, partial $\eta^2 = 0.15$, power = 0.86.

For each group, we conducted one sample Student's *t*-tests to determine if the children learned any of the six tone words at a level better than chance. Our hypothesis was that if the children were using the transitional probability cues within the stream of tones to discover the tone word boundaries, then their ability to correctly recognize a tone word at test should correspond to internal transitional probability of the tone words, with children's performance

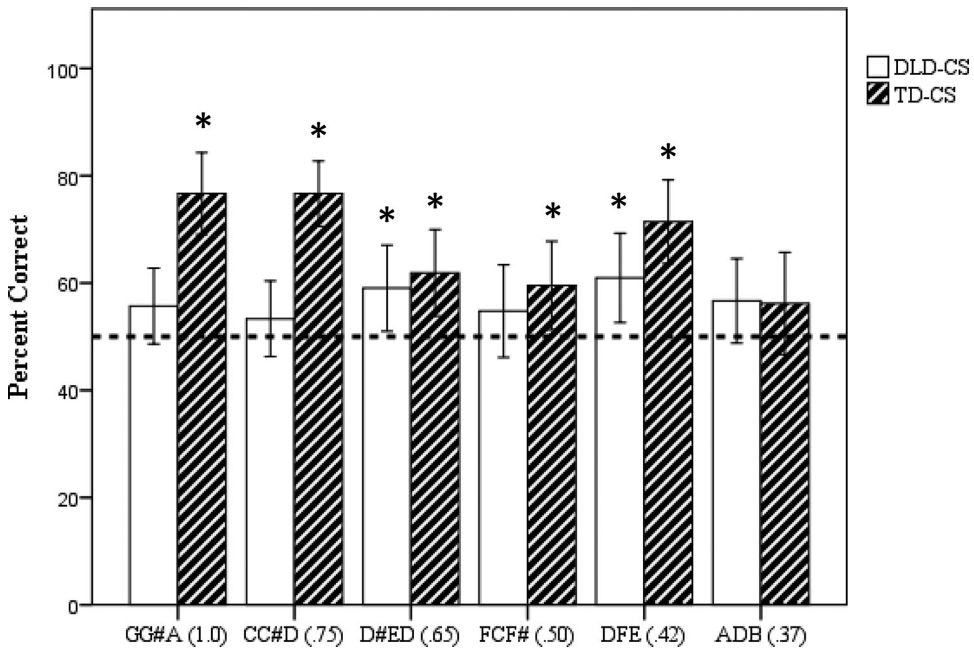

FIGURE 2 Accuracy by individual tone word for the Catalan–Spanish-speaking children with developmental language disorder (DLD-CS) and typically developing controls (TD-CS). *Note*: The internal transitional probability of each word is shown in parentheses. Chance equals 50%. The error bar reflects 95% confidence intervals around the means. *Significantly different from chance on the *t*-test

TABLE 3 Individual Student's *t*-test values versus chance for the six tone words on the statistical word-learning task for the Catalan–Spanish-speaking children with developmental language disorder (DLD) and children with normal language (TD)

	DLD-CS			TD-CS		
	Mean	***t*(35)**	***p*-value**	**Mean**	***t*(35)**	***p*-value**
GG#A (1.0)	55.7	1.64	0.11	76.6	7.10	**< 0.00**
CC#D (0.75)	53.3	0.96	0.34	76.6	8.90	**< 0.00**
D#ED (0.65)	59.0	2.29	**< 0.05**	61.9	3.00	**< 0.00**
FCF# (0.50)	54.7	1.12	0.27	59.5	2.34	**< 0.05**
DFE (0.42)	60.9	2.67	**< 0.01**	71.4	5.58	**< 0.00**
ADB (0.37)	56.6	1.71	0.09	56.1	1.32	0.19

Note: Significance *p*-values is shown in bold.

being greater for the words having the highest transition probability (i.e., GG#A, CC#D) as compared with the tone words having the lowest internal transitional probability (i.e., ADB).

As can be seen in Figure 2 and Table 3, the TD-CS group exhibited the pattern we expected, where the strength of the transitional probability of the tone words mirrored the children's segmentation accuracy, with the children identifying 5:6 tone words having the highest internal transitional probability better than chance, but not the tone word having the lowest transitional probability (ABD (0.37)). In contrast, the children in the DLD-CS group learning only 2:6 tone words better than chance ((D#ED (0.65) and DFE (0.42)). The results of this analysis suggest that the TD-CS controls were able to attend to and use the transition probability cues in the stream of tones to discover the tone-word boundaries. In contrast, despite performance being above chance overall for the children with DLD-CS, the tone words learned by the DLD-CS children were not those having the highest internal transitional probability suggesting that they were not tracking the statistical information in the input.

Relationship between SWL and vocabulary

We used multiple regression analysis to investigate the relationship between SWL and expressive and receptive vocabulary. For the Catalan–Spanish participants, SWL performance was significantly correlated with expressive vocabulary ($r = 0.36$, $p < 0.01$), receptive vocabulary ($r =$

TABLE 4 Regression model to predict Expressive Vocabulary for Catalan–Spanish (CS)-speaking children with developmental language disorder (DLD-CS) and children with normal language (TD-CS)

Model	β-coefficient	R^2	R^2 change	F change
Order 1				
Age (months)	−0.17	0.01	0.01	0.66
NVIQ	−0.10	0.02	0.01	0.68
SWL	0.33**	0.14	0.12**	9.95**
Group	1.22**	0.68	0.53**	107.9**
Group × SWL	−0.58	0.66	0.00	1.21
Order 2				
Age (months)	−0.09	0.01	0.01	0.66
NVIQ	−0.06	0.02	0.01	0.68
Group	0.83**	0.67	0.65**	134.5**
SWL	0.05	0.68	0.00	0.49
Group × SWL	−0.58	0.68	0.00	1.21

Note: Regression model predicting children's statistical word learning (SWL) performance on the Kaufmann Brief Intelligence Test: NVIQ score and the Kaufman Brief Intelligence, Spanish version: Expressive Vocabulary score (Kaufman & Kaufman, 2004).
*$p < 0.05$; **$p < 0.01$.

TABLE 5 Regression model to predict Receptive Vocabulary for Catalan–Spanish (CS)-speaking children with developmental language disorder (DLD-CS) and children with normal language (TD-CS)

Model	β-coefficient	R^2	R^2 change	F change
Order 1				
Age (months)	0.01	0.00	0.00	0.01
NVIQ	0.18	0.03	0.03	2.20
SWL	0.28*	0.11*	0.08*	6.02*
Group	0.76**	0.58**	0.47**	73.1**
Group × SWL	−0.60	0.58	0.00	0.99
Order 2				
Age (months)	0.01	0.00	0.00	0.01
NVIQ	18	0.03	0.03	2.20
Group	0.76**	0.58**	0.55**	87.1**
SWL	−0.02	0.58	0.00	0.00
Group × SWL	−0.60	0.58	0.00	0.99

Note: Regression model predicting children's statistical word learning (SWL) performance on the Kaufmann Brief Intelligence Test: NVIQ score and the Peabody Picture Vocabulary Test (PPVT-III Spanish Version).
*$p < 0.05$; **$p < 0.01$.

0.30, $p < 0.05$) but not age ($r = 0.02$, $p = 0.87$) or NVIQ ($r = 0.05$, $p = 0.69$).

Inspection of the histograms and normal P-P plots of residuals suggested that the analysis described below met the assumptions of linear regression. We considered two models. For the first model, the dependent variable was Expressive vocabulary and for the second model the dependent variable was Receptive vocabulary. For both models, we considered two orders of independent variable entry to inspect independent variances accounts for by SWL and group membership (DLD, TD): (1) age, NVIQ, SWL, group, and group × SWL; and (2) age, NVIQ, group, SWL, and group × SWL. Age and NVIQ were entered first because these two variables were considered control variables.

In the first model, SWL and group were significant predictors of Expressive vocabulary independent of age and NVIQ as indicated by the significant B-coefficient and R^2 change result by adding SWL and group to the model following age and NVIQ (Table 4). Children with DLD had significantly lower Expressive vocabulary scores than TD controls. Critically, those children who were better statistical word learners had better Expressive vocabulary. The pattern was the same in the second model for Receptive vocabulary, where SWL and group were significant predictors of Receptive vocabulary independent of age and NVIQ (Table 5). Again, children with DLD had significantly lower Receptive vocabulary and children with better SWL ability had better Receptive vocabulary.

Use of SWL to rule in/rule out DLD

Likelihood ratio (LH) analyses were conducted using presence/absence of DLD-CS as the gold standard to determine whether Catalan–Spanish-speaking children's ability to track the statistics within the stream of tones might serve as a screening tool to detect and diagnose in children with DLD (Haynes et al., 2006; Sackett et al., 1991). To determine the LH ratio for a positive result based on per cent correct at test on the SWL task (SWL-PC), the true positive rate (proportion of children with DLD-CS with a total SWL-PC at or below x-determined cut-off) was divided by the false-positive rate (proportion of TD-CS children with total SWL-PC at or below x-determined cut-off). The LH ratio for a negative test result sufficient to rule out the presence of DLD was set at a total SWL-PC at x-determined cut-off or higher. The negative test was calculated by dividing the false-negative rate (proportion of children with DLD who had total SWL-PC at or above x-determined cut-off) by the true-negative rate (proportion of TD children with total SWL-PC or above x-determined cut-off). We used Haynes et al.'s (2006) criteria to classify a positive test (i.e., accurately ruling in the disorder) which includes: (1) 'High' as defined as LH ratio of ≥ 20 having a probability of ≥ 95% that the disorder is present; (2) 'Intermediate High' defined as a LH ratio between 1 and 20; and (3) 'Indeterminated' defined as a LH close to 1.0. For a negative test (i.e., accurately ruling out the disorder) Haynes et al. define the LH to be as close to 0 as possible.

In the present study, LH ratios were used to calculate the cut off scores to maximize the ability to 'rule in' and

TABLE 6 Number and proportion of children (negative and positive rates) for different cut-off values based on percentage correct at test on the statistical word-learning task (SWL-PC) developmental disorder diagnosis (DLD) for the Catalan–Spanish (CS)-speaking children with and without DLD

	DLD-CS					TD-CS				
SWL-PC	Test negative ≥		Test positive ≤			Test negative ≥		Test positive ≤		
Cut-off	*n*	Proportion	*n*	Proportion	+LH	*n*	Proportion	*n*	Proportion	–LH
45	**30**	**0.8571**	**5**	**0.1429**	**50.00**	34	0.9714	1	0.0286	0.88
50	21	0.6000	14	0.4000	40.67	32	0.9143	3	0.0857	0.66
55	21	0.6000	14	0.4000	20.80	30	0.8571	5	0.1429	0.70
60	14	0.4000	21	0.6000	10.50	21	0.6000	14	0.4000	0.67
65	8	0.2286	27	0.7714	10.50	17	0.4857	18	0.5143	0.47
70	2	0.0571	33	0.9429	10.65	15	0.4286	20	0.5714	0.13
75	1	0.0286	34	0.9714	10.42	**11**	**0.3143**	**24**	**0.6857**	**0.09**
80	1	0.0286	34	0.9714	10.26	8	0.2286	27	0.7714	0.13
85	1	0.0286	34	0.9714	10.13	5	0.1429	30	0.8571	0.20

Note: For the cut-off of 50% the participants at ≤ 50 were computed for +LH and the participants > 50 were computed for –LH. The best cut-offs for +LH (< 45) and–LH (> 75) are highlighted in bold.

'rule out' DLD, we first calculated the number and proportion of children in the DLD and TD groups whose scores fell above (test negative) or below (test positive) for a given SWL-PC value (Table 6). For each cut-point, we then calculated the positive LH (+LH) and the negative LH (–LH) ratios corresponding to each cut-off to identify Catalan–Spanish-speaking children.

As can be seen in Table 6, the analysis revealed that for the Catalan–Spanish-speaking children, the most discriminating positive test result was a score of ≤ 45%. This was determined by dividing the true positive rate (the number of children with DLD who had total SWL-PC ≤ 45%, which is 5/35 or 0.1429) by the false-negative rate (the number of children with TD with total SWL-PC ≤ 45%, which is 1/35, or 0.0286). The +LH ratio for a score of ≤ 45% (0.1429/0.0286) for ruling a child in to the DLD group was 5.00, indicating that a child who scored ≤ 45% on the SWL tasks was five times more likely to be a Catalan–Spanish-speaking child with DLD as opposed to a child with normal language. Based on Haynes et al. (2006), this positive test result (≤ 45%) would be considered as an 'intermediate high' +LH, indicating that additional diagnostic testing would be required to correctly classify Catalan–Spanish-speaking children as having DLD.

The analysis also revealed that the most discriminating negative test –LH value (i.e., children without DLD) was SWL-PC ≥ 75%. This was determined by dividing the false-negative rate (the number of children with DLD who had total SWL-PC this high, which is 1/35 or 0.0286) by the true-negative rate (the number of children with TD who had total SWL-PC this high, which is 11/35, or 0.3143). The resulting –LH ratio for a score of ≥ 75% was 0.09, indicating that a child whose score on the SWL task was ≥ 75% could be 'ruled out' of the DLD group with a high degree of confidence. LHs for levels of SWL-PC performance between 45% and 75% cut points was 1.26, indicating that a child whose score fell within this range would not be able to be correctly classified either as DLD or TD.

Children with above-chance performance

There was a subset of the Catalan–Spanish-speaking children in both the DLD and TD groups whose SWL performance was greater than chance (> 50%) (DLD $n = 21$; TD $n = 32$). One question was whether the performance for this subset of children with DLD mirrored that of the TD controls. The standardized behavioural assessment scores for this subset of children are shown in Table 7. Similar to the larger cohort, the DLD and TD groups also did not differ in age or NVIQ, and the language scores for the children with DLD were significantly poorer than those of the TD controls.

Results for the children having above-chance performance in the DLD and TD groups are shown in Figure 3. The mean for the subset of children with DLD-CS was 63.7% (SD = 7.3), whereas the mean for the TD-CS group was 68.9% (SD = 12.2). Single-sample *t*-tests (two-tailed) where chance equals 50%, confirmed that the performance of each group was significantly better than chance (DLD-CS $t(20) = 8.6$, $p < 0.001$ and TD-CS $t(32) = 8.7$, $p < 0.001$). An analysis of variance (ANOVA) revealed that the performance of the children with DLD-CS did not differ from that of the TD-CS controls $F(1, 53) = 3.03$, $p = 0.08$, partial $\eta^2 = 0.05$, power = 0.40.

The results by tone word are shown for both groups in Figure 4 and the results of the one sample Student's *t*-tests for the individual words are shown in Table 8. The performance for the above chance TD group mirrored that of the larger cohort of TD-CS controls, with the

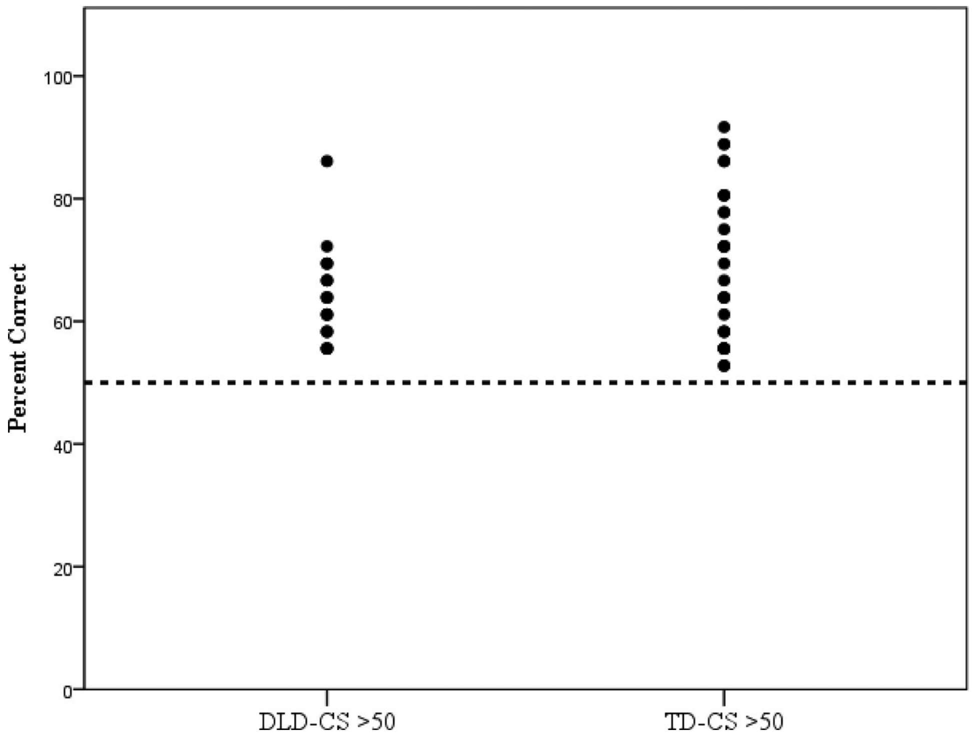

FIGURE 3 Percentage of correct answers for the tone version of the statistical word learning (SWL) task for the group of Catalan-speaking children with developmental language disorder (DLD-CS) and the group of typically developing (TD-CS) children with above chance (i.e., >50%) performance on the SWL task. *Note*: Chance equals 50%

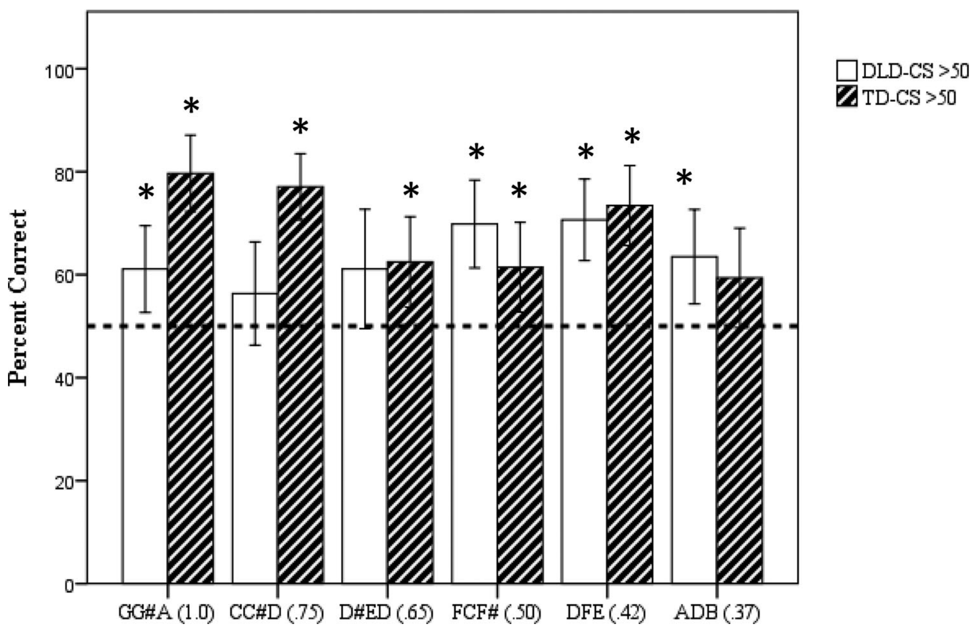

FIGURE 4 Accuracy by individual tone word for the subset of Catalan–Spanish-speaking children with developmental language disorder (DLD-CS) and typically developing controls (TD-CS) who had above chance performance on the statistical word-learning task. *Note*: The internal transitional probability of each word is shown in parentheses. Chance equals 50%. The error bar reflects 95% confidence intervals around the means. *Significantly different from chance on the *t*-test

TABLE 7 Age and standardized scores for language and cognitive assessment measures for Catalan–Spanish (CS)-speaking children with developmental language disorder (DLD-CS) and typically developing (TD-CS) children having above chance (AC) performance on the statistical word-learning task

	DLD-CS ($n = 21$)			TD-CS ($n = 32$)			Comparison	
Variable	**Mean**	**SD**	**Range**	**Mean**	**SD**	**Range**	$t(68)$	p**-value**
Age (months)	104.85	16.06	66–155	108.25	20.18	67–153	−0.65	0.52
K-BIT mat (IQ)[a]	100.14	11.60	82–119	103.78	10.10	88–125	−1.21	0.23
CELF-CLS[b]	72.52	11.30	45–89	108.19	5.82	95–125	−13.34	**< 0.01**
CELF-ELS[c]	73.66	9.56	52–87	107.81	7.81	89–128	−14.24	**< 0.01**
CELF-RLS[d]	75.57	9.43	59–97	105.78	5.76	94–118	−14.49	**< 0.01**
Concepts & Directions[e]	5.66	2.15	1–9	11.69	1.75	8–15	−11.18	**< 0.01**
K-BIT voc[i]	79.95	9.88	63–95	106.81	10.41	83–127	−9.37	**< 0.01**
PPVT-III[g]	78.76	10.39	60–99	105.44	12.60	83–125	−8.06	**< 0.01**

Note: [a]K-BIT mat = Kaufman Brief Intelligence, Spanish version: Non-verbal intelligence score (Kaufman & Kaufman, 2004) Scaled scores ($M = 100$, SD = 15).
[b]CELF-4 CLS = Spanish Clinical Evaluation of Language Fundamentals, 4th Edition: Core Language score (Wiig et al., 2006). Scaled scores ($M = 100$, SD = 15).
[c]CELF-4 ELS = Spanish Clinical Evaluation of Language Fundamentals, 4th Edition: Expressive Language score (Wiig et al., 2006). Scaled scores ($M = 100$, SD = 15).
[d]CELF-4 RLS = Spanish Clinical Evaluation of Language Fundamentals, 4th Edition: Receptive Language score (Wiig et al., 2006). Scaled scores ($M = 100$, SD = 15).
[e]Clinical Evaluation of Language Fundamentals (Wiig et al., 2006): Oral Directions Receptive Subtest Score ($M = 10$, SD = 3).
[i]K-BIT vocabulary. Kaufman Brief Intelligence, Spanish version: Expressive vocabulary score (Kaufman & Kaufman, 2004) Scaled scores ($M = 100$, SD = 15).
[g]PPVT-III = Peabody Picture Vocabulary Test, Third Edition, Spanish version (Dunn, Dunn & Arribas, 2006) Scaled scores ($M = 100$, SD = 15).
Significance of the p-values highlighted in bold.

TABLE 8 Individual Student's t-test values versus chance for the six tone words on the statistical word-learning task for the Catalan–Spanish-speaking children with developmental language disorder (DLD-AC) and typically developing (TD-AC) controls having above-chance performance on the statistical word-learning task (SWL)

	DLD-AC			TD-AC		
	Mean	$t(20)$	p**-value**	**Mean**	$t(31)$	p**-value**
GG#A (1.0)	61.1	2.75	**0.01**	79.6	8.14	**< 0.00**
CC#D (0.75)	56.3	1.32	0.20	77.0	8.59	**< 0.00**
D#ED (0.65)	61.1	2.00	0.06	62.4	2.90	**< 0.00**
FCF# (0.50)	69.8	4.85	**< 0.00**	61.4	2.68	**< 0.01**
DFE (0.42)	70.6	5.43	**< 0.00**	73.4	6.15	**< 0.00**
ADB (0.37)	63.4	3.06	**< 0.00**	59.3	1.98	0.06

Note: Significance p-values is shown in bold.

children again exhibiting the expected pattern of performance where the children learned 5:6 words having the highest internal transitional probability. Surprisingly, although overall performance for the above-chance group of children with DLD did not differ from that of the TD controls, the word-level analysis revealed that the individual tone words learned by the above chance group with DLD-CS differed from the words learned by both groups of TD-CS controls. Specifically, the children with DLD-CS having above chance performance learned GG#A (1.0) and the three words having the lowest internal transitional probability FCF# (0.50), DFE (0.42) and ADB (0.37). This suggests that these children with DLD appeared to be able to use the transitional probability cues in the exposure stimuli to discover the tone word boundary of the tone word having the highest internal transitional probability. However, the DLD group's learning of the words having low internal transitional probability suggests that they may also have been attending to cues other than transitional probability either during exposure to the tone language or at test.

In Experiment 1 we used a tone version of a SWL task to examine SWL in Catalan–Spanish-speaking children with and without DLD and observed that after 21 min of exposure the performance of the children with DLD was significantly poorer than their TD peers. We also observed that children's SWL ability was significantly correlated with expressive and receptive vocabulary. The LH analysis revealed that a score of ≤ 45% on the tone version of the SWL task had a high degree of likelihood of identifying a child as DLD. In Experiment 2 we ask if the pattern of results for an English-speaking cohort of children with and without DLD matched on age and NVIQ mirrors that of the Catalan–Spanish-speaking cohort in Experiment 1.

EXPERIMENT 2

Method

Participants

Data were analysed from a total of 43 children (16 girls and 27 boys), 24 English-speaking children with DLD-E

TABLE 9 Age and standardized scores for language and cognitive assessment measures for English-speaking (E) children with developmental language disorder (DLD-E) and typically developing (TD-E) children

	DLD-E ($n = 24$)			TD-E ($n = 19$)			Comparison	
Variable	**Mean**	**SD**	**Range**	**Mean**	**SD**	**Range**	$t(40)$	p**-value**
Age (months)	110.1	16.5	77–135	106.6	17.9	89–154	1.07	0.29
Leiter-R[a]	98.9	9.3	87–119	102.1	6.1	91–113	1.2	0.21
CELF-3 ELS[b]	70.0	12.2	50–84	107.5	10.6	88–125	10.5	**< 0.00**
CELF-3 RLS[c]	67.6	14.4	50–98					
Concepts & Directions[d]	5.1	2.0	3–10	10.3	2.4	5–14	9.16	**< 0.00**
EVT[e]	81.6	8.4	68–109	102.8	11.7	83–124	6.74	**< 0.00**
PPVT-III[f]	91.7	10.2	69–112	106.5	11.7	84–126	4.52	**< 0.00**

Note: [a]Leiter–R (International Performance Scale—Revised; Roid & Miller, 1997), Age-scaled scores ($M = 100$, SD = 15).
[b]CELF-3 ELS = Clinical Evaluation of Language Fundamentals, Expressive Language Score (Semel et al., 1995), Age-scaled scores ($M = 100$, SD = 15).
[c]CELF-3 RLS = Clinical Evaluation of Language Fundamentals, Receptive Language Score (Semel et al., 1995), Age-scaled scores ($M = 100$, SD = 15).
[d]Clinical Evaluation of Language Fundamentals (Semel et al., 1995): Oral Directions Receptive Subtest Score ($M = 10$, SD = 3).
[e]EVT = Expressive Vocabulary Test (Williams, 1997), Age-scaled scores ($M = 100$, SD = 15).
[f]PPVT-III = Peabody Picture Vocabulary Test, Third Edition (Dunn & Dunn, 1997), Age-scaled scores ($M = 100$, SD = 15).
Significance of the p-values are highlighted in bold.

($M_{\text{Age}} = 9;1$) and 19 chronological age- and NVIQ-matched typically developing children with normal language TD-E ($M_{\text{Age}} = 8;9$). The children in Experiment 2 were part of a larger study examining language and cognitive processing in children with and without DLD conducted between 2000 and 2006 in Madison, Wisconsin.[3] The children were recruited from local public and parochial schools in the greater Madison metropolitan area. All the participating families completed an informed consent form and a background information questionnaire. A summary of the results of all the tests administered to the children was given to the family upon request as a token of gratitude for their commitment and contribution to the study.

All participants all met the following inclusion criteria: (1) normal NVIQ; (2) normal hearing sensitivity at time of testing (American National Standards Institute, 1997); (3) normal or corrected vision; (4) normal oral and speech production as confirmed by a certified Speech Language Pathologist; and (5) from a monolingual English-speaking home environment. Children were excluded if parents reported: (1) neurodevelopmental disorder, (2) emotional or behavioural disturbances, (3) frank neurological signs or (4) seizure disorders or use of medication to control seizures. English was the primary language spoken by the children.

The children in the DLD-E group had a formal diagnosis of language impairment and were receiving speech–language services at the time of the study. Children in the TD-E group were at grade level in school, had no history, or diagnosis, of language-learning disability and had never received speech and language services. To confirm participant's language status, standardized testing was completed by a certified speech–language pathologist at the time of the study and included Leiter International Performance Scale—Revised (Roid & Miller, 1997) and the Clinical Evaluation of Language Fundamentals—3 (Semel et al., 1995), Peabody Picture Vocabulary Test—Third Edition (PPVT-III; Dunn & Dunn, 1997) and the Expressive Vocabulary Test (EVT; Williams, 1997). For the children in the DLD-E group, both Expressive and Receptive Language composite scores on the CELF-3 were ≥ 1.5 SD below age level expectations. For the children in the TD-E group, Expressive Language composite scores, and standard scores on the Concepts & Directions subtest from the Receptive Language portion of the CELF-3 were at or above age level expectations.[4] Similar to the Catalan–Spanish-speaking cohort, non-verbal IQ was within normal limits for the participants, and the two groups did not differ in age or the number of females/males per group $\chi(1) = 0.462$, $p = 0.49$ (Table 9).

Stimuli and procedures

The stimuli and procedures for Experiment 2 were the same as Evans et al. (2009). In Experiment 1 the Catalan–Spanish-speaking children listened to the tone stream tape only for a total of 21 min; however, because the Madison study examined children's learning after double the exposure to the tone language, the English-speaking children in Experiment 2 heard the recording of the tone stream twice, without a break, for 42 continuous minutes. The English-speaking children heard the same digital recording of the 36 test trials in the same order as the children in Experiment 1. To ensure that the children understood the task, children also completed a series of practice trials containing word–non-word pairs derived from words in English (e.g., *com-pu-ter* versus *pu-ter-com*). All the children were able to successfully complete the practice trials and no participant was excluded from the study due an inability to understand the task.

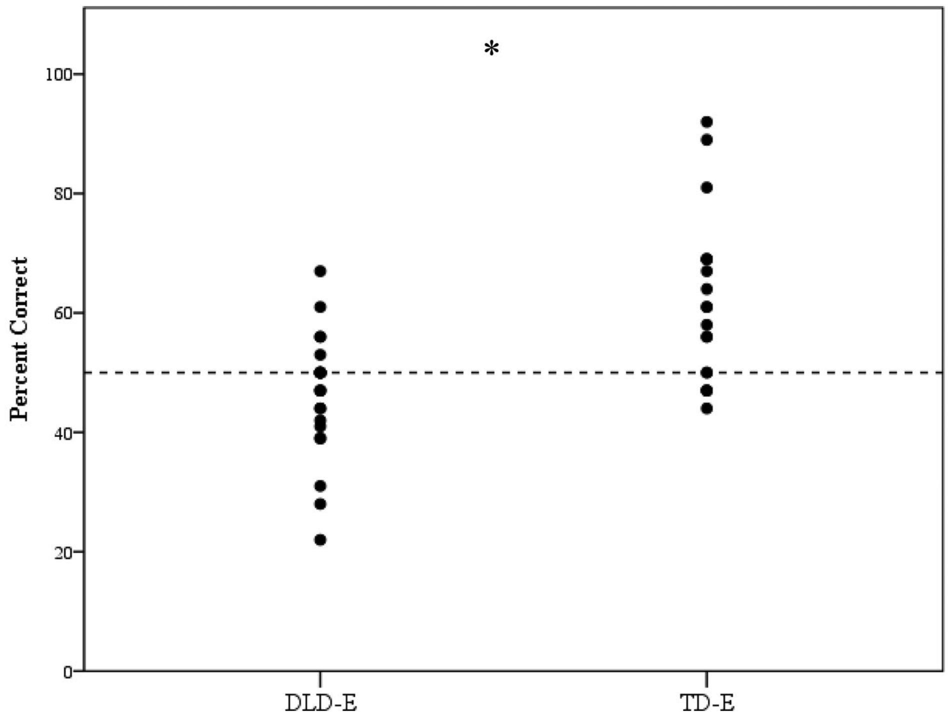

FIGURE 5 Percentage of correct answers for the tone version of the statistical word learning (SWL) task for the group of English-speaking children with developmental language disorder (DLD-E) and the group of typically developing (TD-E) children

Results

Statistical word learning (SWL)

Accuracy for children in the DLD-E and TD-E groups is shown in Figure 5. The mean for the children with DLD-E was 46.6% (SD = 9) and for the TD-E group was 61.9% (SD = 13.9). To be consistent with the analysis in Experiment 1, an ANCOVA with NVIQ as a covariate was conducted. The results revealed that the DLD-E group's performance was significantly poorer than that of the TD group, $F(1, 40) = 16.18$, $p < 0.001$, partial $\eta^2 = 0.28$, power = 0.97. Single-sample *t*-tests (two-tailed) where chance equals 50% indicated that the performance of the DLD-E group did not different from chance $t(23) = -1.65$, $p = 0.11$, whereas the performance was for the TD-E group was significantly better than chance $t(18) = 3.75$, $p < 0.001$. The results by word are shown for the English-speaking cohort in Figure 6 and Table 10. The children in the TD-E group learned 2:6 words greater than chance (CC#D (0.75) and DFE (0.42)), whereas the children with DLD-E did not learn any word significantly *greater* than chance.

TABLE 10 Individual Student's *t*-test values versus chance for the six tone words on the statistical word-learning task for the English-speaking children with developmental language disorder (DLD-E) and children with normal language (TD-E)

	DLD-E			TD-E		
	Mean	***t*(23)**	***p*-value**	**Mean**	***t*(19)**	***p*-value**
GG#A (1.0)	34.0	−4.04	< **0.00***	48.2	−0.34	0.73
CC#D (0.75)	48.6	−0.25	0.79	64.9	2.62	**0.01**
D#ED (0.65)	44.4	−1.35	0.18	61.4	1.78	0.09
FCF# (0.50)	43.7	−1.30	0.20	59.6	1.93	0.06
DFE (0.42)	52.0	0.40	0.68	68.4	3.62	< **0.00**
ADB (0.37)	42.3	−1.66	0.11	51.7	0.23	0.81

Note: *Performance for the DLD group for GG#A ($p = 0.00$) was significantly *below* chance.
Significance *p*-values is shown in bold.

Relationship between SWL and vocabulary

We again considered two models, the first where the dependent variable was Expressive vocabulary, and the second dependent variable was Receptive vocabulary. For the English-speaking cohort, SWL performance was significantly correlated with expressive vocabulary ($r = 0.38$, $p < 0.05$), receptive vocabulary ($r = 0.54$, $p < 0.01$) but not age ($r = -0.08$, $p = 0.69$) or NVIQ ($r = 0.14$, $p = 0.35$). Inspection of the histograms and normal P-P plots of residuals suggested that the analysis described below met the assumptions of linear regression. Same as for the CS cohort, we considered two orders of independent variable entry to inspect independent variances accounts for by SWL and group membership (DLD, TD): (1) age, NVIQ, SWL, group, and group × SWL; and (2) age, NVIQ, group, SWL, and group × SWL. Age and NVIQ were entered first because these two variables were considered control variables.

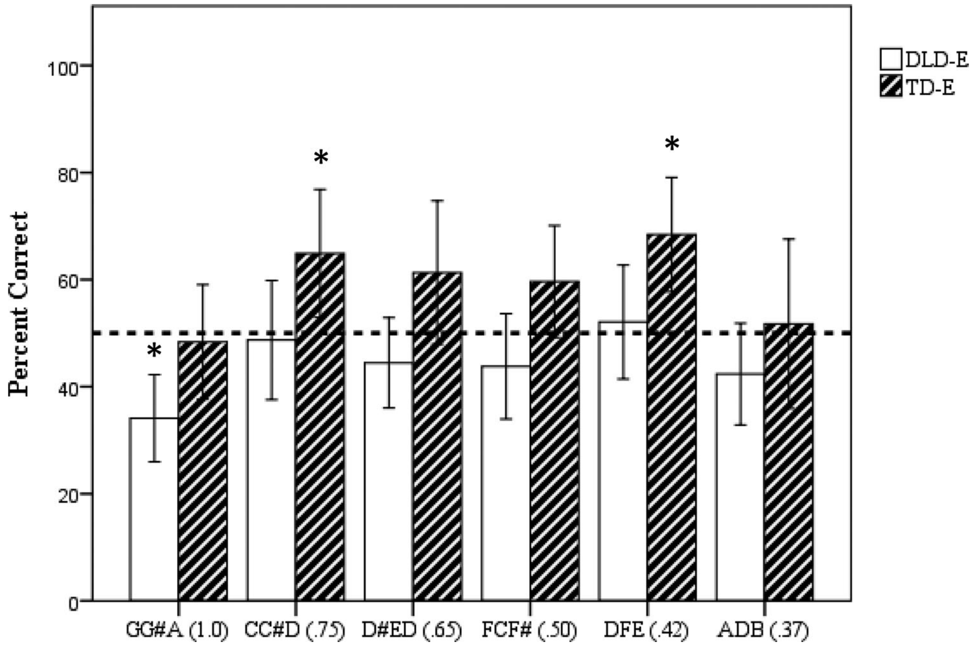

FIGURE 6 Accuracy by individual tone word for the English-speaking children with developmental language disorder (DLD-E) and typically developing controls (TD-E). The internal transitional probability of each word is shown in parentheses. *Note*: Chance equals 50%. The error bar reflects 95% confidence intervals around the means. *Significantly different from chance on the *t*-test

TABLE 11 Regression model to predict Expressive Vocabulary for English (E)-speaking children with developmental language disorder (DLD-E) and children with normal language (TD-E). Regression model predicting children's Statistical Word Learning (SWL) performance on the Leiter-R Nonverbal Intelligence Test (NVIQ) and Expressive Vocabulary Test (EVT)

Model	β-coefficient	R^2	R^2 change	F change
Order 1				
Age (months)	–0.21	0.04	0.04	2.05
NVIQ	0.29	0.13	0.08	3.80
SWL	0.50**	0.37	0.24**	15.2**
Group	0.58**	0.60	0.23**	22.2**
Group × SWL	1.0	0.62	0.01	1.74
Order 2				
Age (months)	–0.21	0.04	0.04	2.05
NVIQ	0.29	0.13	0.08	3.80
Group	0.68**	0.58**	0.45**	42.1**
SWL	0.18	0.60	0.02	2.33
Group × SWL	1.0	0.62	0.01	1.74

Note: $*p < 0.05$; .$**p < 0.01$.

TABLE 12 Regression model to predict Receptive Vocabulary for English (E)-speaking children with developmental language disorder (DLD-E) and children with normal language (TD-E). Regression model predicting children's Statistical Word Learning (SWL) performance on the Leiter-R Nonverbal Intelligence Test (NVIQ) and the Peabody Picture Vocabulary Test (PPVT-III)

Model	β-coefficient	R^2	R^2 change	F change
Order 1				
Age (months)	–0.08	0.00	0.00	0.30
NVIQ	0.40**	0.17**	0.16**	7.91**
SWL	0.33**	0.27*	0.10*	5.82*
Group	0.46**	0.42	0.14**	9.49**
Group × SWL	0.37	0.42	0.00	0.14
Order 2				
Age (months)	–0.08	0.00	0.00	0.30
NVIQ	0.40**	0.17**	0.16**	7.91**
Group	0.50**	0.41**	0.24**	16.5**
SWL	0.08	0.42	0.00	0.32
Group × SWL	0.37	0.42	0.00	0.14

Note: $^{*}p < 0.05$; $^{**}p < 0.01$.

In the first model, SWL and group were significant predictors of Expressive vocabulary independent of age and NVIQ as indicated by the significant *B*-coefficient and R^2 change result by adding SWL and group to the model following age and NVIQ (Table 11). Children with DLD-E had significantly lower Expressive vocabulary scores than TD controls. Critically, those children who were better statistical word learners had better Expressive vocabulary. In the second model for Receptive vocabulary, although NVIQ was a significant predictor of Receptive vocabulary, SWL and group accounted for significant independent variance in Receptive vocabulary independent of age and NVIQ (Table 12). Again, children with DLD-E had significantly lower Receptive vocabulary and English-speaking

TABLE 13 Number and proportion of children (negative and positive rates) for different cut-off values based on percentage correct at test on the statistical word-learning task (SWL-PC) developmental disorder diagnosis (DLD) for the English-speaking children with and without DLD

SWL-PC	DLD-E					TD-E				
	Test negative ≥		Test positive ≤			Test negative ≥		Test positive ≤		
Cut-off	*n*	Proportion	*n*	Proportion	+LH	*n*	Proportion	*n*	Proportion	–LH
45	**15**	**0.6250**	**9**	**0.3750**	**7.13**	**18**	**0.9474**	**1**	**0.0526**	**0.66**
50	5	0.2083	19	0.7917	2.51	13	0.6842	6	0.3158	0.30
55	4	0.1667	20	0.8333	2.64	13	0.6842	6	0.3158	0.24
60	2	0.0833	22	0.9167	1.94	10	0.5263	9	0.4737	0.16
65	1	0.0417	23	0.9583	1.52	7	0.3684	12	0.6316	0.11
70	0	0.0000	24	1.0000	1.19	**3**	**0.1579**	**16**	**0.8421**	**0.00**
75	0	0.0000	24	1.0000	1.19	3	0.1579	16	0.8421	0.00
80	0	0.0000	24	1.0000	1.19	3	0.1579	16	0.8421	0.00
85	0	0.0000	24	1.0000	1.19	3	0.1579	16	0.8421	0.00

Note: For the cut-off of 50% the participants at 50 and below were computed for the +LH and the participants above 50 were computed for the –LH.

TABLE 14 Age and standardized scores for language and cognitive assessment measures for English-speaking (E) children with developmental language disorder (DLD-E) and typically developing (TD-E) children. Subjects that performed above chance in the SWL task

	DLD ($n = 5$)			TD ($n = 13$)			Comparison	
Variable	Mean	SD	Range	Mean	SD	Range	$t(40)$	*p*-values
Age (months)	108.2	10.4	99–123	104.0	17.9	89–154	1.07	0.61
Leiter-R[a]	103.4	4.0	98–108	101.7	6.1	91–113	1.2	0.54
CELF-3 ELS[b]	68.0	15.5	50–84	109.3	10.6	88–125	10.5	**< 0.00**
CELF-3 RLS[c]	73.2	13.8	50–86					
Concepts & Directions[d]	6.8	2.2	4–10	9.7	2.4	5–14	9.16	**< 0.05**
EVT[e]	82.2	9.4	68–93	104.9	11.7	89–124	6.74	**< 0.00**
PPVT-III[f]	92.6	13.3	75–112	106.3	11.7	93–118	4.52	**< 0.05**

Note: [a]Leiter-R (International Performance Scale—Revised; Roid & Miller, 1997), Age-scaled scores ($M = 100$, SD = 15).
[b]CELF-3 ELS = Clinical Evaluation of Language Fundamentals, Expressive Language Score (Semel et al., 1995), Age-scaled scores ($M = 100$, SD = 15).
[c]CELF-3 RLS = Clinical Evaluation of Language Fundamentals, Receptive Language Score (Semel et al., 1995), Age-scaled scores ($M = 100$, SD = 15).
[d]Clinical Evaluation of Language Fundamentals (Semel et al., 1995): Oral Directions Receptive Subtest Score ($M = 10$, SD = 3).
[e]EVT = Expressive Vocabulary Test (Williams, 1997), Age-scaled scores ($M = 100$, SD = 15).
[f]PPVT-III = Peabody Picture Vocabulary Test, Third Edition (Dunn & Dunn, 1997), Age-scaled scores ($M = 100$, SD = 1.
Significance of the *p*-values is shown in bold.

children with better SWL ability had better Receptive vocabulary.

Use of SWL to rule in/rule out DLD

The same LH analysis was conducted to determine if English-speaking children's ability to track the statistics within the stream of tones might serve as a screening tool to detect and diagnose children with DLD in the English-speaking cohort.

LH ratios

As can be seen in Table 13, the true positive rate +LH ratio (0.375/0.052) for the English-speaking cohort also was ≤ 45% for a +LH of 7.13, indicating that an English-speaking child with a score of ≤ 45% on the tone SWL task was seven times more likely to be a child with DLD-E than a child with normal language. Alternatively, most discriminating negative test results were ≥ 70% for the English-speaking cohort, which resulted in a –LH for a negative test result equalled 0.00. Thus, an English-speaking child who had a score ≥ 70% at test on the tone version of the task could be 'ruled out' of the DLD group with an excellent degree of confidence.

Children with above-chance performance

Similar to the Catalan–Spanish cohort, there was a small group of DLD and TD controls whose performance was greater than chance (DLD: $n = 5$; CA: $n = 13$). The standardized behavioural assessment scores for this subset of children in each group are shown in Table 14.

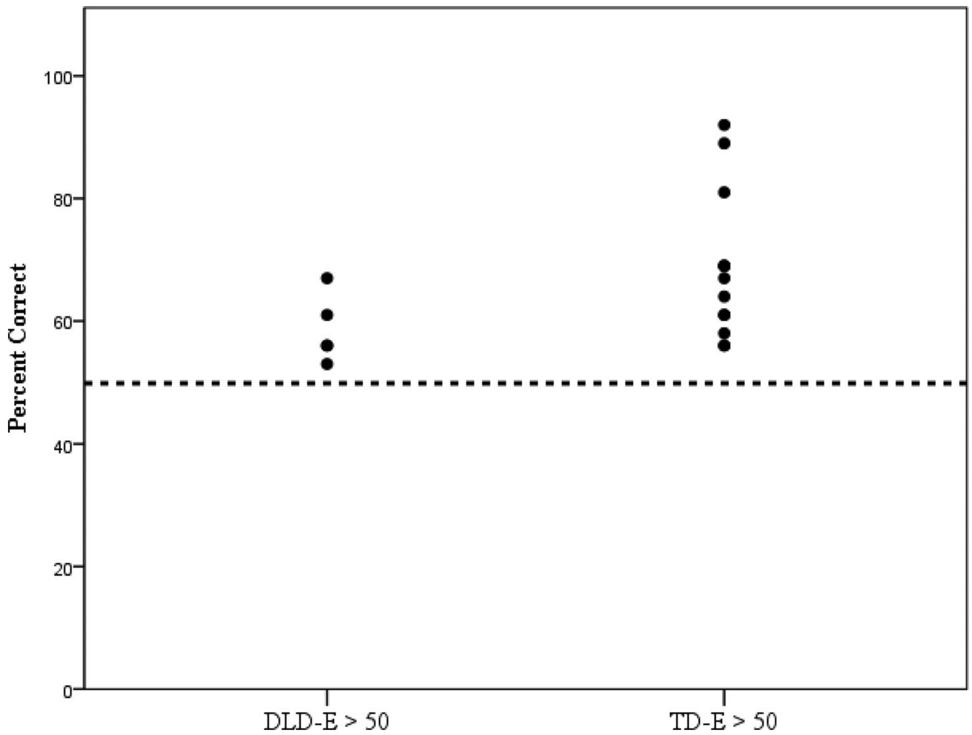

FIGURE 7 Percentage of correct answers for the tone version of the statistical word learning (SWL) task for the group of English-speaking children with developmental language disorder (DLD-E) and the group of typically developing (TD-E) children with above chance (i.e., >50%) in the statistical word-learning task. *Note*: Chance equals 50%

Results for the English-speaking children with above-chance performance are shown in Figure 7. The mean for the subset of children with DLD-E was 58% (SD = 5.0) and for the TD-E controls was 68% (SD = 11). Single-sample *t*-tests (two-tailed) where chance equals 50% confirmed that the performance for both groups was significantly better than chance (DLD-E $t(4) = 3.3$, $p < 0.05$; TD-E $t(32) = 5.6$, $p < 0.001$). An ANOVA revealed that the performance for the children with DLD-E did not differ significantly from that of the above chance TD-E children $F(1, 17) = 3.38$, $p = 0.08$, partial $\eta^2 = 0.17$, power = 0.40.

Performance for each word is shown in Figure 8 and the sample Student's *t*-tests are presented in Table 15. The subset of children with DLD-E with above chance performance learned 1:6 words greater than chance (DFE (0.42)). In contrast, the above chance TD-E group learned 3:6 words greater than chance (CC#D (0.75), D#ED (0.65) and DFE (42)). Unlike the Catalan–Spanish cohort, these findings suggest that children in both the DLD and TD groups were attending to cues other than transitional probability either in the input stream or in the test trials.

TABLE 15 Individual Student's *t*-test values versus chance for the six tone words on the statistical word-learning task for the subset of English-speaking children with developmental language disorder (DLD-E) and children with normal language (TD-E) having above chance performance on the statistical word-learning task

	DLD-E			CA-E		
	Mean	***t*(4)**	***p*-value**	**Mean**	***t*(12)**	***p*-value**
GG#A (1.0)	50.0	0.0	1.0	56.4	1.10	0.29
CC#D (0.75)	56.6	0.40	0.70	71.7	3.15	< **0.00**
D#ED (0.65)	46.6	–1.0	0.37	71.7	3.04	< **0.01**
FCF# (0.50)	46.6	–0.40	0.70	60.2	1.86	0.08
DFE (0.42)	66.6	3.1	< **0.05**	74.3	4.16	< **0.00**
ADB (0.37)	53.3	0.30	0.77	57.6	.82	0.42

Note: Significance *p*-values is shown in bold.

Catalan–Spanish versus English-speaking cohorts

A key difference between the two experiments in the current study was the amount of time the Catalan–Spanish and English-speaking cohorts were exposed to the tone language. To avoid potential ceiling effects, the Catalan–Spanish study was designed to mirror that of Saffran et al. (1999), and the children heard the exposure stimuli for a total of 21 min. However, in the original Evans et al. (2009) study, the children heard the tone word stimuli for a total of 42 min because study asked whether additional exposure time would aid in the DLD group's ability to discover the tone words in the stream of tones. An ANOVA comparing SWL for the Catalan–Spanish and English-speaking TD children revealed no effect of group $F(1, 53) = 1.6$, $p = 0.21$, partial $\eta^2 = 0.03$, power = 0.23; however, a follow-up analysis revealed a significant group

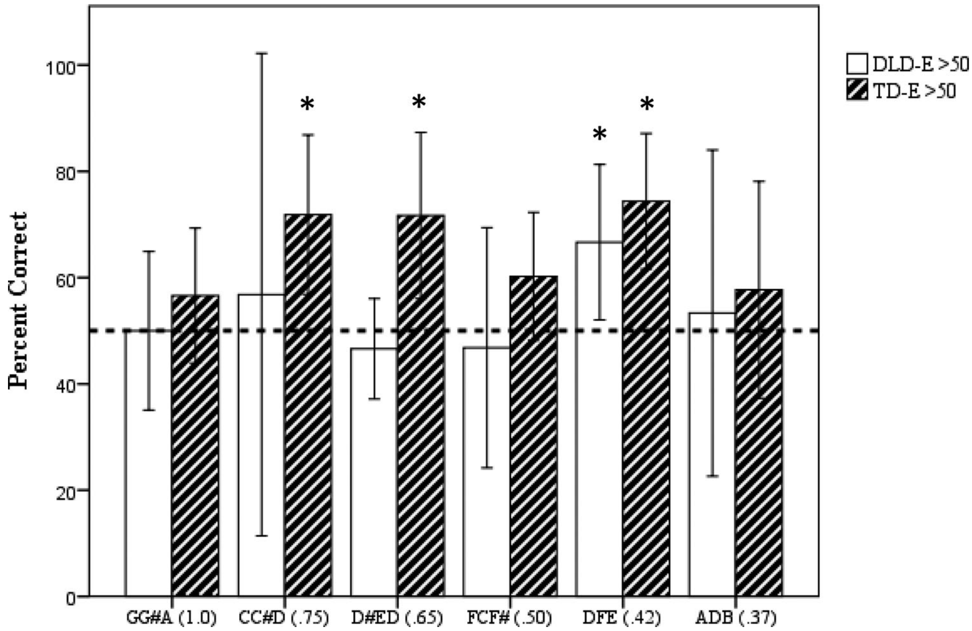

FIGURE 8 Accuracy by individual tone word for the subset of English-speaking children with developmental language disorder (DLD-E) and typically developing controls (TD-E) having above chance performance on the statistical word-learning task. *Note*: The internal transitional probability of each word is shown in parentheses. Chance equals 50%. The error bar reflects 95% confidence intervals around the means. *Significantly different from chance on the *t*-test

by tone word interaction $F(1, 57) = 7.4$, $p < 0.01$, partial $\eta^2 = 0.12$, power = 0.76, where the TD-CS children learned 2:6 words (GG#A (1.0), CC#D (0.75)) significantly better than the TD-E controls. An ANOVA revealed that SWL learning for the Catalan–Spanish-speaking children with DLD was significantly better than the English-speaking children with DLD $F(1, 87) = 15.01$, $p < 0.001$, partial $\eta^2 = 0.20$, power = 0.96. Follow-up analysis revealed that the DLD-CS cohort learned 3:6 words (GG#A (1.0), D#ED (0.65), ADB (0.37)) significantly better than the DLD-E cohort $F(1, 57) = 0.35$ $p = 0.55$, partial $\eta^2 = 0.00$, power = 0.09. These results indicate that for both the DLD and TD children, performance was better after 21 min of exposure indicating that double exposure to the tone stimuli did not result in improved performance for the English-speaking children.

A second question is whether differences in age, NVIQ and language abilities were related to SWL for the Catalan–Spanish and English cohorts. The two TD groups did not differ in age $F(1, 53) = 0.32$ $p = 0.84$, partial $\eta^2 = 0.00$, power = 0.05; NVIQ $F(1, 53) = 0.32$, $p = 0.57$, partial $\eta^2 = 0.00$, power = 0.08, CELF-E $F(1, 53) = 0.12$ $p = 0.72$, partial $\eta^2 = 0.00$, power = 0.06; receptive vocabulary $F(1, 53) = 0.00$ $p = 0.92$, partial $\eta^2 = 0.00$, power = 0.05, or expressive vocabulary $F(1, 53) = 1.34$ $p = 0.25$, partial $\eta^2 = 0.02$, power = 0.20. The Catalan–Spanish and English-speaking children brought different music, cultural and language exposure to the study, which may account for the difference in the individual words learned by the two TD groups.

The two DLD cohorts also did not differ in age $F(1, 58) = 0.87$ $p = 0.35$, partial $\eta^2 = 0.01$, power = 0.15; NVIQ $F(1, 58) = 0.00$ $p = 0.96$, partial $\eta^2 = 0.00$, power = 0.05; CELF-E $F(1, 58) = 1.39$ $p = 0.24$, partial $\eta^2 = 0.02$, power = 0.21; or expressive vocabulary $F(1, 58) = 2.6$ $p = 0.11$, partial $\eta^2 = 0.04$, power = 0.35. In contrast, receptive language abilities of the Catalan–Spanish children with DLD were better than those of the English-speaking children with DLD: CELF-R $F(1, 58) = 9.4$ $p < 0.01$, partial $\eta^2 = 0.14$, power = 0.85, receptive vocabulary $F(1, 58) = 22.1$ $p < 0.001$, partial $\eta^2 = 0.28$, power = 0.99. This suggests that better SWL learning is linked to better receptive language ability in children with DLD.

DISCUSSION AND CONCLUSIONS

The results of this investigation replicated those of prior studies that have reported deficits in SWL in children with DLD. In the present study, performance for both groups of children with DLD was significantly poorer than that of their TD peers. Regression analysis also revealed that for both cohorts, SWL accounted for a significant amount of unique variance in Receptive and Expressive vocabulary, indicating that children with better SWL ability have better vocabulary regardless of their native language. LH ratio analysis revealed that a Catalan–Spanish child with a score of ≤ 45% on the SWL task would be five times more likely to have DLD, and English-speaking child with this score would be seven times more likely to have DLD. These LH values are in line with LH values reported by Ellis Weismer et al. (2000) for non-word repetition tasks, indicating that, similar to non-word repetition, SWL appears to be effective in distinguishing between affected and unaffected children regardless of their native language.

The word-level analysis revealed that the Catalan–Spanish and English-speaking children with DLD appeared to be attending to cues other than transitional probability. The PDH account posits that children with DLD may compensate for procedural memory deficits by relying on declarative memory (Ullman & Pullman, 2015). In this study we used the term *implicit learning* to refer to a form of learning in which the children were able/unable to extract statistical structure from the input stimulus but where the evidence of this learning was manifested in change in their performance, not in their ability to verbalize this knowledge explicitly. Although implicit learning in adults has typically been studies using tasks such as artificial grammar learning and SRT tasks, studies suggest that SWL is another form of implicit learning (e.g., Karuza et al., 2013) and seen as falling under the purview of implicit learning and the basal ganglia, however, there is some evidence of declarative memory/medial temporal lobe (MTL) and hippocampal involvement in statistical learning in typical individuals (Turk-Browne et al., 2009). Further, SWL studies using both syllables and tones in patients having MTL damage and/or complete loss of bilateral hippocampus show at chance SWL as compared with healthy participants (Covington et al., 2018; Schapiro et al., 2014). The hippocampus is able to support the rapid binding of arbitrarily related elements and their temporal relations (Eichenbaum & Cohen, 2001) and the above results suggest that declarative memory may contribute to unconscious processing of relational binding on a rapid time course characteristic of statistical learning as well.

In the present study, the children with DLD may have been relying on declarative memory as a compensatory strategy both during the exposure to the language and at test. Notably, the tone language used in the study was designed by Saffran et al. (1999) to ask: Does the ability to implicitly track transitional probability within a stream of speech extend to non-linguistic tone stimuli, not if a child is relying on declarative memory as a compensatory learning strategy? However, the word level analysis of TD controls with above-chance performance in both the Catalan–Spanish and English cohorts suggests that in addition to transitional probability cues, there may have been other cues inherent in the tone stream and/or the test items themselves. It has been argued that fragments and/or chunks provide good coding of information (Perruchet & Pacton, 2006). A chunking strategy representation is based on the participants' memorization of fragments of strings and reflects reliance on the declarative memory system. Although not designed specifically to address the question, in the current study the children with DLD may have been using a chunk-based strategy reflecting a reliance on declarative memory as a compensatory strategy both in processing the tone stimuli, and in an attempt to manage the 2AFC test format which may have placed extensive demands on already poor phonological working memory in the children with DLD (Ellis Weismer et al., 2000).

In this study, the absolute performance on the SWL task differed for the Catalan–Spanish children with DLD from that of the English children with DLD. Moreover, the Catalan–Spanish cohort was bilingual and the English cohort was monolingual, and the exposure time differed across the two cohorts. That the same relationship between SWL and vocabulary and LH cut points was observed across these two distinct cohorts of children is an intriguing finding. Recently there has been increased interest in examining the consistency in children's performance on various implicit learning tasks (Arnon, 2020; West et al., 2018). Although the results of these studies have been somewhat equivocal, the LH values from this project suggest that performance on auditory SWL tasks such as the one used in the current study may be sensitive to individual differences in children's language abilities in children regardless of their clinical status.

Another unique aspect of the current study was that the children in the Catalan–Spanish cohort were simultaneous bilinguals. Demographically, the rapid growth in the number of children being raised in bilingual language learning environments poses a diagnostic dilemma for researchers, educators and practitioners due to potential overlap in the performance of child second language (L2) acquisition and DLD on standardized tests. Studies of bilingual children with DLD (BIDLD) aim at disentangling the effects of bilingualism from those of DLD, making use of both models of bilingualism and models of language impairment (Armon-Lotem, 2012). To date, studies examining performance on tasks designed to tap implicit learning skills in BIDLD is sparse, however, Park et al. (2018) observed poor performance on procedural sequential learning tasks in bilingual school-age children with and without DLD that mirrors those reported by Lum et al. (2014) supporting the proposal that DLD may be characterized by failure of learning on SRT tasks. What is notable about Park et al. (2018) was that the children in their study had various language backgrounds (Korean, Chinese, German, Bengali, French and Spanish) in addition to speaking English. The findings from the current study suggest that deficits in SWL are evident in children with DLD regardless of the child's native language. Further, the findings from the current study suggest that whether the child is bi- or monolingual, the tone version of the SWL task may serve as an additional cross-linguistic screening tool to detect and diagnose children with DLD.

Consistent with prior work by Graf et al (2007), this study shows that SWL appears to be a fundamental learning mechanism that children rely on to acquire vocabulary and that children with poor SWL abilities may experience significant challenges in their vocabulary acquisition. Clinically, the results from this study show that poor performance on auditory SWL tasks where children are required to track transitional probability cues within a stream of input may be indicative of the presence of language deficits, regardless of the child's native language, and whether or not the child is bilingual or monolingual. Clinically, similar to non-word repetition tasks (Ellis Weismer et al., 2000), although the results from this study suggest that poor performance on these types of SWL tasks is highly suggestive of DLD, but may not be diagnostically sufficient for ruling in and ruling out language impairment in children.

ORCID

Nadia Ahufinger https://orcid.org/0000-0003-2029-4475
Julia L. Evans https://orcid.org/0000-0002-0863-4150

NOTES

[1] In Catalonia, both Spanish and Catalan are official languages and therefore proficiency in both languages is, if not native, native-like. Accordingly, it is very difficult to find a monolingual child. Bilingualism in Catalonia is considered as balanced because almost all the population is proficiencient in two languages such that their skills in each language match those of a native speaker of the same age. The vehicular language at schools is Catalan, but Spanish is also taught at school. Although Catalan is an institutional and an official language, Spanish has a big presence in the social life of the country. Everyone has the right to use both languages in all social contexts.

[2] All children were classified as either TD or DLD based on standard clinical practice in Catalonia–Spain, which is based on the Spanish version of the standardized test where all stimuli are presented in Spanish, but if children answered correctly in Catalan they were given credit for their answer.

[3] A subset of this larger group of children with and without DLD (DLD = 15; TD = 15) was reported by Evans et al. (2009).

[4] Although two children in the TD-E group had scores that were < 1 SD on the Concepts and Directions subtest of the CELF-3, they were included in the TD-E control group because their performance in school was at grade level and their ELS on the CELF-3 as well as their PPPVT and EVT scores were at or above age-level expectations.

ACKNOWLEDGEMENTS

We thank Lisbeth Simon-Heilmann, Karin Phillips and Kristen Ryan for their assistance with this research. Especially, we thank the children and their families who generously contributed their time. We thank the Associació del Trastorn Específic del Llenguatge de Catalunya (ATELCA) and all the schools that participated in the study. Especially, we thank the participation of the Escola Santa Anna (Premià de Dalt), the Centres de Recursos per a Deficients Auditius de Catalunya (CREDA) and the Equips d'Assessorament Psicopedagògic de Catalunya (EAP). We also thank the teachers, speech–language pathologists and other professionals who contributed to the study. Finally, Coral Pérez and Joan Tarrida are thanked because they contributed to the data collection in Catalonia.

CONFLICTS OF INTEREST

There are no conflicts of interest.

FUNDING

This work was supported by the Ministerio de Economía y Competitividad [grant number 2016EDU2016-75368-P and BES-2014-070511]; the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) de la Generalitat de Catalunya [grant number 2017SGR387]; the National Institute on Deafness and Other Communication Disorders (NIDCD) [grant number R0105263] awarded to Dr Evans; and the NICHD [grant number P30HD03352], awarded to the Waisman Center.

REFERENCES

American National Standards Institute. (1997) American National Standard: methods for Calculation of the Speech Intelligibility Index. Acoustical Society of America.

Armon-Lotem, S. (2012) Introduction: bilingual children with SLI–the nature of the problem. *Bilingualism: Language and Cognition*, 15(1), 1–4.

Arnon, I. (2020) Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods*, 52(1), 68–81.

Bishop, D.V.M. (2014) Ten questions about terminology for children with unexplained language problems. *International Journal of Language & Communication Disorders*, 49(4), 381–415.

Catts, H.W., Bridges, M.S., Little, T.D. & Tomblin, J.B. (2008) Reading achievement growth in children with language impairments. *Journal of Speech, Language, and Hearing Research*, 51 (6), 1569–1579.

Conti-Ramsden, G., Mok, P.L., Pickles, A. & Durkin, K. (2013) Adolescents with a history of specific language impairment (SLI): strengths and difficulties in social, emotional and behavioral functioning. *Research in Developmental Disabilities*, 34(11), 4161–4169.

Covington, N.V., Brown-Schmidt, S. & Duff, M.C. (2018) The necessity of the hippocampus for statistical learning. *Journal of Cognitive Neuroscience*, 30(5), 680–697.

Dunn, L.M. & Dunn, L.M. (1997) *Peabody picture vocabulary test.* PPVT-III. Circle Pines, MN: American Guidance Service.

Dunn, L.M., Dunn, L.M. & Arribas, D. (2006) Test de vocabulario en imágenes. PEABODY: PPVT-III. Barcelona: TEA Ediciones.

Durkin, K. & Conti-Ramsden, G. (2007) Language, social behavior, and the quality of friendships in adolescents with and without a history of specific language impairment. *Child Development*, 78(5), 1441–1457.

Eichenbaum, H. & Cohen, N.J. (2001) *From conditioning to conscious recollection: memory systems of the brain*. New York: Oxford University Press.

Ellis Weismer, E, Tomblin, J B., Zhang, X., Buckwalker, P., Hynoweth, J.G. et al. (2000) Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research*, 43(4), 865–878.

Evans, J.L., Saffran, J.R. & Robe-Torres, K. (2009) Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52(2), 321–335.

Graf Estes, K., Evans, J.L., Alibali, M.W. & Saffran, J.R. (2007) Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 18(3), 254–260.

Haebig, E., Saffran, J.R. & Ellis Weismer, S. (2017) Statistical word learning in children with autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry*, 58(11), 1251–1263.

Haynes, R.B., Sackett, D.L., Guyatt, G.H. & Tugwell, P. (2006) How to do clinical practice research: a new book and a new series in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology*, 59(9), 873–875.

Karuza, E.A., Newport, E.L., Aslin, R.N., Starling, S.J., Tivarus, M.E. & Bavelier, D. (2013) The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain and Language*, 127(1), 46–54.

Kaufman, A.S. & Kaufman, N.L. (2004) *Test Breve de Inteligencia Kaufman. K-BIT*. Spanish adaptation: Cordero-Pando, A. and Calonge-Romano, I. Madrid: Pearson.

Kemény, F. & Lukács, Á. (2010) Impaired procedural learning in language impairment: results from probabilistic categorization. *Journal of Clinical and Experimental Neuropsychology*, 32(3), 249–258.

Lammertink, I., Boersma, P., Wijnen, F. & Rispens, J. (2017) Statistical learning in specific language impairment: a meta-analysis. *Journal of Speech, Language, and Hearing Research*, 60(12), 3474–3486.

Lammertink, I., Boersma, P., Wijnen, F. & Rispens, J. (2020) Children with developmental language disorder have an auditory verbal statistical learning deficit: evidence from an online measure. *Language Learning*, 70(1), 137–178.

Lum, J.A., Conti-Ramsden, G., Morgan, A.T. & Ullman, M.T. (2014) Procedural learning deficits in specific language impairment (SLI): a meta-analysis of serial reaction time task performance. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior*, 51, 1–10.

Lum, J.A., Conti-Ramsden, G., Page, D. & Ullman, M.T. (2012) Working, declarative and procedural memory in specific language impairment. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior*, 48(9), 1138–1154.

Lum, J.A. & Conti-Ramsden, G. (2013) Long-term memory: a review and meta-analysis of studies of declarative and procedural memory in specific language impairment. *Topics in language disorders*, 33(4), 282–297.

Lum, J.A., Gelgic, C. & Conti-Ramsden, G. (2010) Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders*, 45(1), 96–107.

Mayor-Dubois, C., Zesiger, P., Van Der Linden, M. & Roulet-Perez, E. (2012) Nondeclarative learning in children with specific language impairment: predicting regularities in the visuomotor, phonological, and cognitive domains. *Child Neuropsychology*, 20(1), 14–22.

Mishkin, M., Malamut, B. & Bachevalier, J. (1984) Memories and habits: two neural systems. *Neurobiology of learning and memory*, 65–77.

Obeid, R., Brooks, P.J., Powers, K.L., Gillespie-Lynch, K. & Lum, J.A. (2016) Statistical learning in specific language impairment and autism spectrum disorder: a meta-analysis. *Frontiers in Psychology*, 7, 1245.

Park, J., Miller, C.A., Rosenbaum, D.A., Sanjeevan, T., Van Hell, J.G., Weiss, D.J., et al;, (2018) Bilingualism and procedural learning in typically developing children and children with language impairment. *Journal of Speech, Language, and Hearing Research*, 61(3), 634–644.

Perruchet, P. & Pacton, S. (2006) Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, 10(5), 233–238.

Plante, E., Gomez, R. & Gerken, L. (2002) Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders*, 35(5), 453–462.

Poldrack, R.A., Prabhakaran, V., Seger, C.A. & Gabrieli, J.D. (1999) Striatal activation during acquisition of a cognitive skill. *Neuropsychology*, 13(4), 564–574.

Roid, G.H. & Miller, L.J. (1997) *Leiter international performance scale—revised*. Leiter-R. Wood Dale, IL: Stoelting.

Sackett, D.L., Hayes, R.B., Guyatt, G.H. & Tugwell, P. (1991) *Clinical epidemiology: a basic science for clinical medicine*, second edition. New York: Little, Brown.

Saffran, J.R. & Estes, K.G. (2006) Mapping sound to meaning: connections between learning about sounds and learning about words. *Advances in Child Development and Behavior*, 34, 1–38.

Saffran, J.R., Johnson, E.K., Aslin, R.N. & Newport, E.L. (1999) Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52.

Semel, E.M., Wiig, E.H. & Secord, A.W. (1995) *Clinical evaluation of language fundamentals*. CELF-3. San Antonio, TX: Psychological Corporation

Schapiro, A.C., Gregory, E., Landau, B., Mccloskey, M. & Turk-Browne, N.B. (2014) The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience*, 26(8), 1736–1747.

Squire, L.R. (1992) Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, 4(3), 232–243.

Squire, L.R. (1994) Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. In: Schacter, D.L. & Tulving, E. (Eds.) *Memory systems 1994.* Cambridge, Massachusetts: The MIT Press, pp. 203–231.

Squire, L.R. & Knowlton, B.J. (2000) The medial temporal lobe, the hippocampus, and the memory systems of the brain. *The New Cognitive Neurosciences*, 2, 756–776.

Tomblin, J.B., Freese, P.R. & Records, N.L. (1992) Diagnosing specific language impairment in adults for the purpose of pedigree analysis. *Journal of Speech, Language, and Hearing Research*, 35(4), 832–843.

Tomblin, J.B., Mainela-Arnold, E. & Zhang, X. (2007) Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development*, 3(4), 269–293.

Tomblin, J.B., Records, N.L., Buckwalter, P., Zhang, X., Smith, E. & O'brien, M. (1997) Prevalence of specific language impairment in kindergarten children. *Journal of Speech Language Hearing Research*, 40(6), 1245–1260.

Turk-Browne, N.B., Scholl, B.J., Chun, M.M. & Johnson, M.K. (2009) Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *Journal of cognitive neuroscience*, 21(10), 1934–1945.

Ullman, M.T. & Pullman, M.Y. (2015) A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews*, 51, 205–222

Ullman, M.T. & Pierpont, E.I. (2005) Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior*, 41(3), 399–433.

Ullman, M.T., Earle, F.S., Walenski, M., AND & Janacsek, K. (2020) The neurocognition of developmental disorders of language. *Annual Review of Psychology*, 71, 389–417.

West, G., Vadillo, M.A., Shanks, D.R. & Hulme, C. (2018) The procedural learning deficit hypothesis of language learning disorders: we see some problems. *Developmental science*, 21(2), e12552.

Wiig, E.H., Semel, E.M. & Secord, A.W. (2006) *Clinical evaluation of language fundamentals 4: Spanish edition. CELF-4-Spanish*. Madrid: Pearson.

Williams, K.T. (1997) *Expressive vocabulary test*. second edition. EVT-2. Circle Pines, MN: American Guidance Service.

How to cite this article: Ahufinger, N., Ferinu, L., Sanz-Torrent, M., Andreu, L. & Evans, J. L. (2021) Statistical word learning in Catalan–Spanish and English-speaking children with and without developmental language disorder. *International Journal of Language & Communication Disorders*, 1–21. https://doi.org/10.1111/1460-6984.12673