

**Supplemental Material S3.** Discriminative values (sensitivity and specificity).

Study	DLD criteria	Threshold	Sensitivity	Specificity
Conti-Ramsden (2003)	DLD recruited via SLTs or specialist teachers in language. Evaluation through a standardized language test.	<i>p</i> .25	66%	100%
		<i>p</i> .16	59%	100%
Conti-Ramsden & Hesketh (2003)	identical to Conti-Ramsden (2003)	<i>p</i> .25 <i>p</i> .16	66% 59%	85% 85%
Gray (2003)	DLD recruited via SLTs, scoring more than -1.5 <i>SD</i> on two standardized language tests.	na	Time 1: 95% Time 2: 91% Time 3: 95%	Time 1: 100% Time 2: 95% Time 3: 86%
Thal & al. (2005)	Late talkers identified at 16 months as having limited vocabulary (below the 10 percentile) based on parental reports.	na	40%	86%
Bortolini & al., (2006)	DLD diagnosed and followed by a SLT, scoring at or below -1.5 <i>SD</i> on more than one subtest, with and MLU between 2.1 and 4.8.	na	82%	82%
Oetting & Cleveland (2006)	Children referred to SLT services or identified by the teacher as having low language skills, scoring below -1 <i>SD</i> in two standardized tests.	na	56%	92%
Oetting & al. (2008)	Children scoring below the 15th percentile on the NWR and SR tasks, as well as below -1 <i>SD</i> on a standardized test.	70%	56%	92%
Archibald & Joanisse (2009)	Children scoring below the 15th percentile on the NWR and SR tasks, as well as below -1 <i>SD</i> on a standardized test.	<i>p</i> .15	40-46%	60-63%
Deevy & al. (2010)	DLD recruited via SLTs, scoring below -3.25 <i>SD</i> on a standardized test.	> 66% > 68%	PPC method 1: 86% PPC method 2: 79%	PPC method 1: 91% PPC method 2: 89%
Gutiérrez-Clellen & Simon-Cerejido (2010)	Assessment protocol included evidence of clinical concern, 20% of ungrammatical utterances and low performance on monolingual and	70%	English NWRT: 55% Spanish NWRT: 61%	English NWRT: 82% Spanish NWRT: 82%

	bilingual standardized tests (< 60% correct responses).			
Thorardottir & al. (2011)	DLD recruited by SLTs who provided a diagnosis and data from a parental report.	-1 <i>SD</i> -1.28 <i>SD</i> -2 <i>SD</i> 82%/-0.88 <i>SD</i>	85% 85% 77% 92%	86% 88% 93% 86%
Dispaldro & al. (2013)	DLD diagnosed by SLT, score below the norm on standardized tests, MLU analysis, score below -2 <i>SD</i> on third person object clitic pronoun use task	65% 93%	PIC: 100% PPC: 94%	PIC: 100% PIC: 94%
Guiberson & Rodríguez (2013)	Identification of the DLD by a bilingual SLT, parental concerns about language development, language expression scores more than -1.5 <i>SD</i> on a standardized test.	na	PIC: 71% PPC: 48%	PIC: 74% PPC: 78%
Kapalková & al. (2013)	Children with a DLD had a documented history of language impairment since the age of 2–3 years, with a diagnosis based both on clinical judgement and a qualitative assessment of children's language (no standardized tests available in Slovak).	8 items correct	PPC: 94%	PPC: 100%
Stokes & al., (2013)	Diagnosis of DLD by SLT and scores -1.2 <i>SD</i> on a standardized test.	na	77%	97%
Thordardottir & Brandeker (2013)	DLD children diagnosed by SLT, scores below the mean on a French assessment, clinical history.	82%	monolinguals: 92% bilinguals: 85%	monolinguals: 100% bilinguals: 79%
Tuller & al. (2013)	Children with a DLD followed by an SLT, assessment in both languages and use of deviation criteria by Thordardottir (2015).	60%	90%	89,5%

Topbaş & al. (2014)	DLD diagnosed by SLTs and scores on a standardized test more than -1 <i>SD</i> .	-2 <i>SD</i>	89%	87%
Boerma & al. (2015)	DLD diagnosed by SLT, global score of -2 <i>SD</i> or lower on a Dutch assessment battery with score of -1.5 <i>SD</i> or lower on at least two of the four subtests.	77.7	monolinguals: quasi-universal LITMUS: 83%	monolinguals: quasi-universal LITMUS: 90%
		72.7	Dutch LITMUS: 93%	Dutch LITMUS: 93%
		78.1	bilinguals: quasi-universal LITMUS: 83%,	bilinguals: 93% (for both LITMUS tasks)
		63.8	Dutch LITMUS: 63% RNM	
		78.1	monolinguals+bilinguals: quasi-universal LITMUS: 83%	monolinguals+bilinguals: quasi-universal LITMUS: 92%
		72.7	Dutch LITMUS: 87%	Dutch LITMUS: 77%
Ferré & Dos Santos (2015)	DLD children recruited and followed by SLTs, parental report to determine risk factors, standardized assessment in the child's mother tongue and in French. with deviation criteria by Thordardottir (2015).	-2 <i>SD</i>	monolinguals: 90%	monolinguals: 92%
		-1 <i>SD</i>	bilinguals: 73%	bilinguals: 82%
		-1.28 <i>SD</i>	68%	89%
		-1.65 <i>SD</i>	59%	91%
		-2 <i>SD</i>	59%	95%
Guiberson & Rodríguez (2015)	SLT diagnosed language delay, parental concern about language development, score of -1.5 <i>SD</i> or more on a task from a Spanish standardized test.	na	80%	62%
Armon-Lotem & Meir (2016)	Monolingual children with a DLD: diagnosis by SLT, parental report (e.g., parental concern), scores -1.25 <i>SD</i> or more based on monolingual norms on Russian or Hebrew tests. Bilingual children with a DLD: diagnosed by SLT or history of disorder reported by parents/teachers, scores	.71	monolinguals: Russian NWRT: 86%	monolinguals: Russian NWRT: 90%
			Hebrew NWRT: 93%	Hebrew NWRT: 66%
		.71	bilingual children with monolingual norms: Russian NWRT: 44%	bilingual children with monolingual norms: Russian NWRT: 86%
		.86	Hebrew NWRT: 96%	Hebrew NWRT: 48%

	-1.25 <i>SD</i> or more based on bilingual norms on Russian and Hebrew tests.		bilingual children with monolingual norms: Russian NWRT: 70% Hebrew NWRT: 81% Both combined: 65%	bilingual children with monolingual norms: Russian NWRT: 76% Hebrew NWRT: 89% Both combined: 94%
		.79 .71		
De Almeida & al. (2016)	DLD children were followed by SLTs, standardized tests in each language, with deviation criteria by Thordardottir (2015).	80%	monolinguals: 88% bilinguals: 81%	monolinguals: 83% bilinguals: 79%
De Almeida & al. (2017)	identical to De Almeida & al. (2016)	80%	monolinguals: 88% bilinguals: 81%	monolinguals: 83% bilinguals: 79%
Hodges & al., (2017)	Below 15th percentile on a parental vocabulary report, within normal limits for receptive language (> 16th percentile on a standardized test).	60% 51%	MITT: 92% TENR: 90%	MITT: 77% TENR: 87%
Kazemi & Saeednia (2017)	DLD children recruited via SLTs, clinical judgment by two SLTs (no standardized test in Iran).	34: PSC 12.5: PIC	95% 90%	98% 96%
Le Clercq & al. (2017)	DLD children recruited from special education schools, scoring below -1.5 <i>SD</i> in at least two language domains, or at least -2 <i>SD</i> below the general mean measured by Dutch language standardized tests.	11.4% 78.6%	short task: 83% complete task: 87%	short task: 95% complete task: 88%
Dos Santos & Ferré (2018)	DLD children recruited and followed by SLTs, parental report to determine risk factors, standardized assessment in the child's mother tongue and in French with deviation criteria by Thordardottir (2015).	-2 <i>SD</i> -1 <i>SD</i> -1.28 <i>SD</i> -1.65 <i>SD</i> -2 <i>SD</i>	monolinguals: 90% bilinguals: 73% 68% 59% 59%	monolinguals: 92% bilinguals: 82% 89% 91% 95%

Li'el & al. (2018)	DLD children diagnosed and followed by SLTs or on the waiting list. Parental questionnaire (ALDeQ, Paradis & al., 2010) used to assess risk factors.	$\leq 7$ correct items	89%	81%
Tuller & al. (2018)	DLD are followed by SLTs. Assessments in both languages: scores below the mean in two language domains of the dominant language, using the deviation criteria by Thordardottir (2015).	77.5 59.9	French LITMUS: 88% German LITMUS: 92%	French LITMUS: 92% German LITMUS: 90%
McDonald & Oetting (2019)	DLD children scored below -1 <i>SD</i> on an English syntax and lexical subtest.	.65 .76	53% 77%	98% 74%
Pham & Ebert (2020)	DLD children scored below -1 <i>SD</i> on at least one of the parent or teacher concern measures and on at least three of the six direct language measures, and in at least two of the three language domains that were measured (vocabulary, grammar, and narratives).	0.20 0.56 0.85	PIC: 40% PSC: 80% PPC: 90%	95% 79% 79%
Bonifacci & al. (2020)	DLD children were diagnosed by SLTs.	na	50%	91%

*Note:*  $N = 33$ , listed by order of publication date. Abbreviations: PIC = percentage items correct, PPC = percentage phonemes correct, PSC = percentage syllables correct; SLT = speech and language therapist; DLD: children with a development language disorder; SD = standard deviation; MLU = mean length utterance; SR = sentence repetition.

For the inclusion criteria for DLD children, performance is indicated with an arrhythmia sign (-) indicating below-average scores.