

Supplemental Material S1. Two tables are included that provide the diagnostic accuracy metrics for all language sample measures and age groups and the reference measures used across all reviewed studies.

Table S1. Diagnostic accuracy of language sample analysis measures.

Measure	Age	Task	Sensitivity	Specificity	Overall
Mainstream English speakers					
Morphosyntax: Accuracy					
DSS Sentence Point ^a	4;0–4;11	Play	93%	94%	—
	5;0–5;10	Play	100%	100%	—
Errors per C-unit ^b	6	Narrative tell	91%	82%	85%
	8	Narrative tell	94%	80%	84%
Finite Verb Morphology Composite ^{a,b,c,d,e,f,g,h}	3;0–3;11 ^c	Play (50 utterances)	67%	100	83%
	3;0–3;11 ^c	Play (100 utterances)	83%	89%	86%
	3;0–3;11 ^c (SPELT < 87)	Play (100 utterances)	100%	89%	94%
	3;0–3;6 ^c	Play (100 utterances)	100%	75%	88%
	3;6–3;11 ^c	Play (100 utterances)	70%	100%	85%
	3;7–5;9 ^d	Conversation Picture description	84%	100%	—
	4;0–4;11 ^a	Play	93%	94%	—
	5;0–5;10 ^a	Play	91%	93%	—
	4;0–4;6 ^e	Play	100%	100%	—
	5;0–5;6 ^e	Play	92%	93%	—
	5;5–9;9 ^f	Conversation Expository	50%	86%	68%
	5;11–6;3 ^g (MLU < −1 SD)	Play	26	90	—
	5;11–6;3 ^g (PPVT-R < −1 SD)	Play	35	92	—
	5;11–6;3 ^g (NWR < −1 SD)	Play	30	91	—
	4;0–4;11 ^h	Narrative tell	92%	94%	94%
	5;0–5;11 ^h	Narrative tell	100%	90%	92%
	6;0–6;11 ^{b,h}	Narrative tell	82%	90%	89%
	7;0–7;11 ^h	Narrative tell	85%	86%	86%
	8;0–8;11 ^{b,h}	Narrative tell	76%	80%	79%
	9;0–9;11 ^h	Narrative tell	80%	76%	77%
Grammaticality & Utterance Length Instrument ⁱ	4;0–6;11	Narrative retell	83%	92%	—
Noun Morphology Composite ^{d,f}	3;7–5;9 ^d	Conversation Picture description	79%	100%	—
	5;5–9;9 ^f	Conversation Expository	54%	86%	70%

Measure	Age	Task	Sensitivity	Specificity	Overall
Percent Grammatical Utterances/C-units ^{b,j,k}	3;0–3;11 ^j	Picture description	100%	88%	—
	4;0–4;11 ^k	Narrative tell	83%	96%	94%
	5;0–5;11 ^k	Narrative tell	100%	82%	86%
	6;0–6;11 ^{b,k}	Narrative tell	82%	90%	89%
	7;0–7;11 ^k	Narrative tell	92%	88%	89%
	8;0–8;11 ^{b,k}	Narrative tell	88%	84%	85%
	9;0–9;11 ^k	Narrative tell	90%	90%	90%
Percent Sentence Point ^l	3;0–3;11	Picture description	100%	82%	—
Proportion ‘restricted’ utterances ^l	8–10	Narrative tell	83%	88%	—
Percent Verb Tense Usage ^j	3;0–3;11	Picture description	100%	82%	—
Morphosyntax: Proficiency					
DSS Tota ^a	4;0–4;11	Play	79%	94%	—
	5;0–5;10	Play	72%	87%	—
Mean tense/agreement ^a	4;0–4;11	Play	79%	81%	—
	5;0–5;10	Play	64%	80%	—
Mean Top 5 tense/agreement ^a	4;0–4;11	Play	71%	69%	—
	5;0–5;10	Play	73%	87%	—
Tense Marker Total ^c	4;0–4;6	Play	83%	87%	—
	5;0–5;6	Play	77%	80%	—
Tense/Agreement Productivity Score ^{c,e}	3;0–3;5 ^c	Play (100 utterances)	88%	88%	88%
	3;6–3;11 ^c	Play (100 utterances)	90%	80%	85%
	3;0–3;11 ^c	Play (50 utterances)	94	50	72
	3;0–3;11 ^c	Play (100 utterances)	89%	78%	83
	3;0–3;11 ^c SPELT < 87	Play (100 utterances)	100%	100%	100%
	4;0–4;6 ^e	Play	67%	87%	—
	5;0–5;6 ^e	Play	80%	80%	—
Morphosyntax: Length					
Clauses per Sentence ^m	3;0–7;11	Conversation	83%	91%	—
MLU (morphemes) ^{d,f}	3;7–5;9 ^d	Conversation	95%	89%	—
		Picture description			
	5;5–9;9 ^f	Conversation Expository	72%	80%	76%
MLU (SUGAR) ^m	3;0–7;11	Conversation	86%	86%	—
Words per Sentence ^m	3;0–7;11	Conversation	83%	84%	—
Semantics					
Moving Average Type-Token Ratio ⁿ	4–9	Narrative tell	26%	88%	—
Number of Different Words (200w) ⁿ	4–9	Narrative tell	20%	90%	—
Number of Different Words (41u) ⁿ	4–9	Narrative tell	42%	91%	—
Number of Different Words (nar) ⁿ	4–9	Narrative tell	34%	92%	—

Measure	Age	Task	Sensitivity	Specificity	Overall
Total Number of Words ^m	3;0–7;0	Conversation	86%	85%	—
Pragmatics/Discourse					
Story Grammar ^o	4;0–9;11	Narrative retell	70%	84%	81%
TNW + Turns + References + Expansions ^p	8;4–13;2	Expository (Map task)	75%	60%	—
Composite Models					
10 SALT measures ^q	3;0–13;6	Conversation	69%	84%	—
	3;0–5;11	Conversation	87%	87%	—
	6;0–9;11	Conversation	80%	85%	—
	10;0–13;6	Conversation	77%	82%	—
Cohesive ties + % grammatical T-units + subordinate clauses/T-unit + words/subordinate clause ^r	7;6–10;6	Narrative retell	—	—	98%
	8;6–12;6	Narrative retell	—	—	79%
	9;0–11;4	Narrative retell	—	—	83%
MLU + Clauses Per Sentence ^m	3;0–7;11	Conversation	97%	82%	—
MLU + lexical diversity <i>D</i> + age ^s	2;0–4;0	Play	86%	91%	—
MLU-m + NDW + IPSyn total ^t	5;5–6;7	Conversation	—	—	74%
MLU + % structural errors + age ^u	2;6–6;11	Play	81%	83%	—
Noun Composite + MLU ^{d,f}	5;5–9;9 ^f	Conversation Expository	72%	84%	78%
	3;7–5;9 ^d	Conversation Picture description	89%	100%	—
Noun Composite + Verb Composite ^{d,f}	5;5–9;9 ^f	Conversation Expository	62%	86%	74%
	3;7–5;9 ^d	Conversation Picture description	84%	100%	—
Noun Composite + Verb Composite + MLU ^{d,f}	5;5–9;9 ^f	Conversation Expository	72%	88%	80%
	3;7–5;9 ^d	Conversation Picture description	89%	95%	—
Unmarked Verbs + Verb Types ^v	3;4–6;11	Play	89%	90%	—
		Picture description Narrative retell			
Verb Composite + MLU ^{d,f}	5;5–9;9 ^f	Conversation Expository	74%	84%	79%
	3;7–5;9 ^d	Conversation Picture description	95%	95%	—
VP errors + Stage 1 Utterances + Age + 3 – element NP ^w	2;0–4;2	Play	91%	92%	92%
African American English (AAE) and Southern White English (SWE) speakers					
Morphosyntax: Accuracy					
Nonmainstream patterns ^x					
Full model (35 patterns)	4–6	Play	87%	94%	90%
Reduced model (4 patterns)	4–6	Play	74%	90%	84%

Measure	Age	Task	Sensitivity	Specificity	Overall
SWE-specific model (5 patterns)	4–6	Play	87%	95%	—
SAAE-specific model (3 patterns)	4–6	Play	75%	82%	—
Tense & Agreement Forms ^y					
Unmodified Scoring	5	Play	70%	64%	67%
Modified Scoring	5	Play	72%	74%	73%
Strategic Scoring	5	Play	43%	64%	54%
Past Tense (Strategic)	5	Play	70%	85%	77%
Past Tense (Strategic): SWE	5	Play	89%	89%	89%
Past Tense (Strategic): AAE	5	Play	83%	77%	80%
Morphosyntax: Proficiency					
DSS Total ^z	< 6	Play	63%	100%	—
IPSyn Total ^z	< 6	Play	45%	NR	—
Bilingual English Speakers					
Spanish/English Bilinguals					
MLU + Grammaticality + Number of Different Words + % mazes ^{aa}	5;3–8	Narrative retell	—	—	83%
Cantonese/English Bilinguals					
MLU + IPSyn + Lexical diversity ^{ab}	3;8–5;11	Conversation play	78%	78%	78%

Note. DSS = Developmental Sentence Scoring; IPSyn = Index of Productive Syntax; MLU = mean length of utterance; SUGAR = Sampling Utterances and Grammatical Analysis Revised; TNW = total number of words; SALT = Systematic Analysis of Language Transcripts; IPSyn = Index of Productive Syntax; NDW = number of different words; VP = verb phrase. NP = noun phrase; SWE = Southern White English; SAAE = Southern African American English; NR = not reported. ^aSouto et al., 2014. ^bGuo & Schneider, 2016. ^cGuo & Eisenberg, 2014. ^dBedore & Leonard, 1998. ^eGladfelter & Leonard, 2013. ^fMoyle et al., 2011. ^gRudolph et al., 2019. ^hGuo et al., 2020. ⁱCastilla-Earls & Fulcher-Rood, 2018. ^jEisenberg & Guo, 2013. ^kGuo et al., 2019. ^lHoffman, 2009. ^mPavelko & Owens, 2019. ⁿCharest et al., 2020. ^oSchneider et al., 2006. ^pScheffel, 1997. ^qHeilmann et al., 2010. ^rLiles et al., 1995. ^sKlee et al., 2007. ^tHewitt et al., 2005. ^uDunn et al., 1996. ^vFletcher & Peters, 1984. ^wGavin et al., 1993. ^xOetting & McDonald, 2001. ^yOetting et al., 2021. ^zOverton et al., 2021. ^{aa}Smyk, 2012. ^{ab}Ooi & Wong, 2012.

Table S2. Reference measures used in reviewed studies.

Reference measure	Study
Clinical criterion (diagnosis by an SLP/currently receiving treatment)	Charest et al., 2020 ^b Dunn et al., 1996 Gladfelter & Leonard, 2013 ^a Fletcher & Peters, 1984 ^b Gavin et al., 1993 ^b Guo & Eisenberg, 2014 ^b Guo & Schneider, 2016 ^b Guo et al., 2019 ^b Guo et al., 2020 ^b Hoffman, 2009 ^b Klee et al., 2007 ^b Pavelko & Owens, 2019 ^b Heilmann et al., 2010 Liles et al., 1995 Moyle et al., 2011 ^a Oetting & McDonald, 2001 Ooi & Wong, 2012 Overton et al., 2021 Schneider et al., 2006 ^a Souto et al., 2014 ^a
Clinical Evaluation of Language Fundamentals, 3rd Edition (CELF-3)	Charest et al., 2020 Guo & Schneider, 2016 Guo et al., 2019 Guo et al., 2020 Hoffman, 2009 Schneider et al., 2006 ^a
Clinical Evaluation of Language Fundamentals, 4th Edition Spanish (CELF-4 Spanish)	Smyk, 2012
Clinical Evaluation of Language Fundamentals, Preschool 2nd Edition (CELF-P2)	Charest et al., 2020 Guo et al., 2019 Guo et al., 2020 Schneider et al., 2006
Diagnostic Evaluation of Language Variation - Norm Referenced (DELV)	Oetting et al., 2021
EpiSLI System (Test of Language Development, Primary + narrative task)	Hewitt et al., 2005
Preschool Language Scales, 3rd Edition (PLS-3)	Bedore & Leonard, 1998
Peabody Picture Vocabulary Test, Revised (PPVT-R)	Moyle et al., 2011 Bedore & Leonard, 1998 Gavin et al., 1993 Oetting & McDonald, 2001 Rudolph et al., 2019
Reynell Developmental Language Scales (Expressive)	Fletcher & Peters, 1984
Spanish-English Language Proficiency Scales (SELPS)	Smyk, 2012
Sequenced Inventory of Communication Development, Revised (SICD-R)	Gavin et al., 1993 Klee et al., 2007
Structured Photographic Expressive Language Test, 3rd Edition (SPELT-3)	Castilla-Earls & Fulcher-Rood, 2018 Smyk, 2012
Structured Photographic Expressive Language Test, Preschool 2nd Edition (SPELT-P2)	Eisenberg & Guo, 2013 ^a Gladfelter et al., 2013 Souto et al., 2014 Guo & Eisenberg, 2014
Stephens Oral Language Screening Test	Fletcher & Peters, 1984
Test for Auditory Comprehension of Language, Revised (TACL-R)	Moyle et al., 2011

Test for Examining Expressive Morphology (TEEM)	Bedore & Leonard, 1998
Test of Language Development, Primary (TOLD-P)	Pavelko & Owens, 2019
Test of Language Development, Primary 2nd Edition (TOLD-P2)	Bedore & Leonard, 1998
Grammaticality	Oetting & McDonald, 2001
Mean Length of Utterance (MLU)	Smyk, 2012
	Rudolph et al., 2019
	Oetting & McDonald, 2001
Non-Word Repetition (NWR)	Rudolph et al., 2019

Note. ^aReference measure included clinical criterion in addition to other measures. ^bClinical criterion was confirmed by standardized measures.