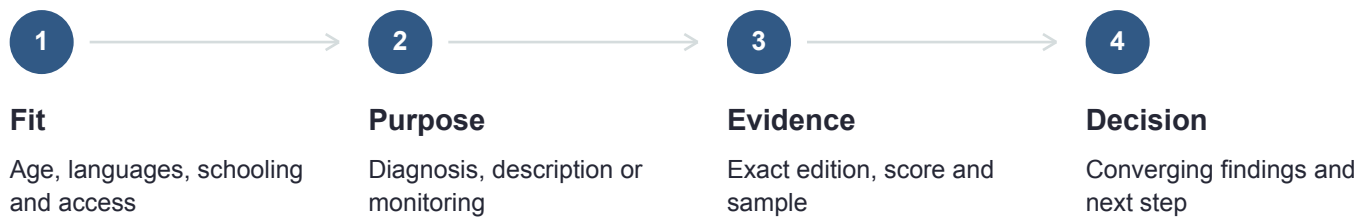


Spanish English assessment comparison manual

Named tools, population fit and diagnostic accuracy you can trace to a source

A practical comparison of named tests, structured samples and learning-based methods for Spanish-English children.



If you need to...	Open this section
Compare language tests and accuracy	Pages 2–8
Compare speech and vocabulary tools	Pages 9–10
Choose narrative or learning measures	Page 11
Interpret accuracy in context	Page 12
Plan a case and record the evidence	Pages 13–14
Check sources and evidence levels	Pages 15–18

Read every accuracy pair with its conditions

Sensitivity is the proportion of children with the target disorder correctly identified. Specificity is the proportion without it correctly classified. Both depend on the sample, reference standard, score and cutoff. They are not a child's probability of having a disorder.

Scope: Selected pediatric resources, checked September 27, 2026. Figures are labeled as primary publisher data, secondary manual reports, or study results. “Not verified” means the accessed sources did not establish the value. It does not mean zero accuracy. Authorized manuals remain necessary for administration and scoring.

For an English-speaking SLP: Plan Spanish access with a qualified bilingual SLP and/or trained interpreter, consistent with each tool's requirements. Clinical interpretation remains the SLP's responsibility. See the companion Working with Interpreters manual and desk checklist. [1, 25]

COMPARE THE POPULATION

Language tests at a glance

Ages use years;months. A Spanish edition can have a very different population from another Spanish edition.

Tool and age	Population and normative sample	Potential use
BESA 4;0–6;11	Spanish-English profiles; N=756 in manual audit. Publisher lists 17 Spanish and 7 English varieties. [2, 3]	Language-specific semantics, morphosyntax and phonology
PLS-5 Spanish Birth–7;11	US/Puerto Rico; Spanish-speaking monolingual and bilingual children. N=1,150. [5]	Early receptive/expressive profile; prescribed dual-language procedure
PLS-5 English US Birth–7;11	English-primary norm group. N=1,400 US children reported by publisher. [7, 8]	English developmental profile when norms and procedure fit
CELF Preschool-2 Spanish 3;0–6;11	Publisher norm sample: more than 450 monolingual/bilingual Spanish-speaking children. [28]	Spanish preschool language structure and content
CELF-4 Spanish US 5;0–21;11	US Spanish-speaking monolingual/bilingual group. Normative N not verified here. [9]	Spanish structural-language information
CELF-5 English US 5;0–21;11	N=2,380 from 47 states; 20% Hispanic is not a bilingual subgroup count. [13]	English language structure and content
CELF-5 Spain 5;0–15;11	Spain norm group; Spanish first language at home. N=922 in 11 age groups. [11, 36]	Population-matched Spanish assessment

The edition check that prevents an avoidable error

CELF-5 Spain is not a new US Spanish norm set. Pearson identifies CELF-4 Spanish as its US-normed Spanish CELF edition. Confirm the exact title, country and manual on the kit. [9, 11]

SAMPLE-SPECIFIC LANGUAGE EVIDENCE

Research studies with bilingual children

These figures answer different questions in different samples. They are not a league table.

Measure and scoring rule	Sensitivity	Specificity
Experimental BESA semantics Both languages below study cutoffs: EN 87.81 and ES 85.02 [4]	93.3%	96.8%
CELF-4 Spanish Core Language cutoff 85 [10]	About 93–94%	65%
CELF-4 Spanish Study-derived cutoff 78 [10]	86%	80%

BESA study: N=78, including 15 with language impairment, ages 4;0–6;11; 40–60% exposure/use in each language; Texas, California and Pennsylvania. This 2008 experimental semantics result does not validate every current BESA score. Apply current manual rules. [4]

CELF-4 Spanish study: Low-income, Spanish-dominant children in Phoenix-area English-only schools; N=656 overall. Accuracy groups: 28 with language impairment and 265 typical children. Reference classification combined Spanish morphology/samples and English evidence. Sensitivity at 85 is 94% in the abstract, 93% in Results. Cutoff 78 had no independent confirmation. [10]

Read the source level: the BESA study includes test developers; the CELF-4 study is independent of its publisher. Neither sample represents every bilingual profile. More manual-reported Spanish estimates are on page 5.

Before using a number

Check age, languages, reference diagnosis, severity spectrum, clinical and typical group sizes, confidence intervals, and whether the cutoff was tested in a new sample. A high point estimate cannot repair poor population fit.

REPORTED MANUAL FIGURES

PLS-5: keep the Spanish and English estimates separate

These are secondary reports of diagnostic accuracy, not new validation in every bilingual population.

Edition and reported group	Sensitivity	Specificity	Cutoff status
PLS-5 Spanish [29–31]	85%	88%	Conflict: –1 SD versus –1.5 SD in accessed reports
PLS-5 English US: older language-disorder group [32]	83%	80%	–1 SD / SS85; exact inequality follows manual
PLS-5 English US: younger developmental-language-delay group [32]	91%	78%	–1 SD / SS85; exact inequality follows manual

PLS-5 Spanish: verify the cutoff before applying it

Bilingualistics pairs 85%/88% with –1 SD; King et al. pairs it with –1.5 SD. The full current technical table was not directly verified. This guide reports the pair but does not resolve the clinical cutoff. The ≤ 77 criterion in the review describes another test used to define the clinical group. [29–31]

Spanish sample: norms N=1,150 US/Puerto Rico Spanish speakers. The review’s accuracy groups span 1;0–7;11 and include moderate/severe disorders. Exact diagnostic denominators are unavailable. The separate tryout sample is not the accuracy sample. [5, 29]

English groups: the review reports clinical groups of 229 (older) and 23 (younger), not verified complete-case denominators. It gives inconsistent lower age bounds: ≥ 3 versus 3;11 for older; 1;0 versus birth for younger, both ending at 7;11 and 3;11 respectively. Do not use these labels as age-specific scoring directions. [32]

Reference and limits: English clinical groups had low scores on another unspecified language measure and received language services. Comparison counts, accuracy confidence intervals, and bilingual subgroup accuracy were not established in these accessed reports. Screening versions and country adaptations have separate evidence. [29, 32]

DO NOT MERGE THESE ROWS

CELF Spanish: manual reports and independent evidence

The same test can perform differently when the sample, reference classification or cutoff changes.

Edition and evidence source	Cutoff / procedure	Sensitivity	Specificity
CELF Preschool-2 Spanish: manual reported secondarily [30, 33]	SS85	86%	89%
CELF Preschool-2 Spanish: independent study reported secondarily [33, 34]	Spanish-only, manual-directed; exact rule unverified	85.7%	64.3%
CELF-4 Spanish: secondary manual review [35]	-1 SD	96%	87%
Same manual report [35]	-1.5 SD	86%	95%
Same manual report [35]	-2 SD	52%	100%

Preschool-2 Spanish: ages 3;0–6;11; publisher norms >450. Clinical N and reference standard behind the manual pair were not verified. The independent study enrolled 30 Mexican American Spanish-English learners ages 4–5; CELF-specific diagnostic denominators were not available. Its rates were retrieved through Castilla-Earls et al., not the full study results. [28, 33, 34]

CELF-4 Spanish: review reports 116 clinical participants ages 5–21 in the US, Puerto Rico and Mexico, plus 116 matched norm controls. Clinical classification used a score below -1.5 SD on an unspecified other test. The review does not reproduce the precise index-combination rule. These are different data from the independent school study on page 3. [35]

CELF-5 Spain is a separate edition

Norms N=922, ages 5;0–15;11. A public diagnostic-accuracy table was not verified. Do not borrow CELF-4 Spanish US or CELF-5 English percentages. No dialect-specific accuracy or accuracy confidence intervals were verified for the rows above. [11, 36]

PRIMARY PUBLISHER FIGURES

English CELF editions: the score rule matters

English norming does not establish diagnostic accuracy for every Spanish-English bilingual child.

Edition and score scope	Reported cutoff	Sensitivity	Specificity
CELF-5 English US: Core + Receptive + Expressive combination [12]	SS80	97%	97%
CELF-5 English US: Core Language alone [12]	-1 SD	99%	90%
CELF Preschool-3 English: table covers Core, Receptive or Expressive scores [37]	-1 SD	96%	70%
Same publisher table [37]	-1.5 SD	89%	87%
Same publisher table [37]	-2 SD	71%	97%

CELF-5 English: ages 5;0–21;11; norms N=2,380. Examiner’s Manual p. 153 reports both pairs. The excerpt does not specify the exact AND/OR rule for the three-score combination. It advises against Core alone at -1.5 or -2 SD. Consult the Technical Manual; do not invent the combination rule. [12, 13]

CELF Preschool-3 English: ages 3;0–6;11; norms N=700 across 41 states. The technical summary labels the accuracy table “Core Language, Receptive Language, or Expressive Language Scores”; it does not establish a separate pair for every index or describe the full classification algorithm. [37]

What these public summaries do not supply

Analytic clinical/control counts, the complete reference-classification procedure, accuracy confidence intervals and Spanish-English subgroup estimates are not verified here. Hispanic representation in norms is not a bilingual language-exposure count. These English figures cannot be transferred to CELF Preschool-2 Spanish.

TILLS: check the language-learning history first

Ages 6;0–18;11. The target condition includes language/literacy disorders, not only developmental language disorder.

Age band	Identification Core sum	Sensitivity	Specificity	Clinical / controls
6;0–7;11	Below 24	84%	84%	51 / 352
8;0–11;11	Below 34	88%	85%	160 / 461
12;0–18;11	Below 42	86%	90%	68 / 442

Primary source: technical manual, Tables 2.1–2.2. Identification Core is the sum of required age-specific subtest standard scores, not the Total score. At or above the cutoff is classified unaffected in this analysis. Use the manual for the correct subtests and scoring. [38]

Who was represented?

Native English speakers included simultaneous bilinguals exposed to English from birth. Sequential English learners were excluded. The publisher advises against diagnostic use for sequential English learners and against translating items. No separate Spanish-English accuracy estimate is supplied. [39]

Keep two counts separate: norm N=1,262; accuracy analysis used 1,255 controls and 279 participants with prior language/literacy diagnoses. Prior diagnosis with service eligibility supplied classification. Subtests and cutoffs were selected using these data; this excerpt does not describe an independent confirmation sample. Accuracy confidence intervals were not verified. [38, 42]

QUICK APPLICATION

Two bilingual children, different fit

Two 9-year-olds use Spanish and English. One has heard English from birth; the other began sustained English exposure at school.

For the first child, examine all other fit and interpretation requirements. For the second, these diagnostic norms are not supported by the publisher's stated eligibility. Gather appropriate evidence across languages.

Clinical move: The label bilingual is not enough to establish norm fit. Language history changes the assessment choice. [1, 39]

CASL-2: compare the original sample and cutoff

English assessment, ages 3;0–21;11. These figures concern the General Language Ability Index (GLAI).

GLAI cutoff	Sensitivity	Specificity
70	41%	99%
75	47%	96%
80	64%	91%
85	74%	84%
90	86%	76%

Source: WPS brief, Table 2, reproducing the original manual analysis. Clinical N=271 with mixed diagnoses, including language disorder, hearing loss, autism, social-pragmatic disorder, intellectual disability, learning disability and developmental delay. Original analytic control N and accuracy confidence intervals were not verified. This is not a DLD-only study. [40]

Norms and bilingual fit: N=2,394, including some bilingual/multilingual participants judged proficient in English. No Spanish-English accuracy cell is supplied. English proficiency and the assessment question still matter; these estimates do not validate translated items. [41]

Do not quietly switch samples

The same WPS brief also reports higher accuracy after restricting the clinical sample to moderate/severe presentations. That reanalysis is a different population and is not independent validation. The table above preserves the original diverse sample. [40]

QUICK APPLICATION

A higher cutoff changes two kinds of error

At cutoff 85, the reported pair is 74%/84%. At cutoff 90, it is 86%/76%.

Sensitivity rises while specificity falls: more affected participants are identified, and more unaffected participants test positive. Neither pair is this child's probability of disorder.

Clinical move: Do not choose a cutoff by the larger percentage alone. Follow the authorized interpretation procedure and use converging evidence.

Speech tools and regional representation

Single-word results need connected speech, intelligibility and language-specific phonological analysis.

Tool or BAPA scoring group	Cutoff	Sensitivity	Specificity
BAPA Spanish/bilingual Spanish [14]	SS 85	94%	90%
Same group [14]	SS 77	72%	92%
Same group [14]	SS 70	56%	95%
BAPA bilingual English [14]	SS 85	92%	87%
Same group [14]	SS 77	92%	93%
Same group [14]	SS 70	77%	96%
GFTA-3 Spanish Sonidos-en-palabras [15]	SS 85	91%	88%

BAPA | 3;0–10;11: N=438 Central Texas children. Spanish speakers were 95% Mexican Spanish; 90% of bilingual/Spanish groups attended dual-language programs. Accuracy analyses added 86 students, 23 in bilingual/Spanish groups. Exact diagnostic denominators and confidence intervals need verification. [14]

GFTA-3 Spanish | 2;0–21;11: US/Puerto Rico Spanish speakers. The public summary does not establish the accuracy-analysis N, reference standard or confidence intervals and contains inconsistent demographic labels. Confirm them in the Spanish manual. [15, 27]

Do not silently repair a publisher table

BAPA's public English-only rows show a non-monotonic pattern across stricter cutoffs. Those estimates are withheld here pending clarification. The Spanish and bilingual-English rows above are transcribed as published. [14]

Vocabulary tools answer a narrower question

Vocabulary can inform an assessment. It cannot stand in for the whole language system.

Tool	Population and sample	Use and boundary
EOWPVT-4 Spanish-Bilingual 2;0–70+ [16]	US Spanish-English bilingual population; publisher reports 1,200+ across child/adult ages.	Picture naming in either language under prescribed rules. Not comprehensive language or language proficiency.
ROWPVT-4 Spanish-Bilingual 2;0–70+ [17]	Co-normed bilingual edition; 1,200+ population. Child/dialect cells not verified here.	Receptive picture vocabulary. Do not substitute English fifth-edition materials or norms.
English-Spanish Vocabulary Inventory [18]	Toddler research samples N=52, 13 and 22; mostly Mexican-heritage families in western US communities.	Caregiver description across languages. Not normed; no validated DLD cutoff supplied.

Accuracy status: No DLD sensitivity/specificity pair was verified for the two one-word publisher pages. ESWI authors identify further norming and diagnostic work as needed. Reliability, correlation with another measure and a large norm sample are different evidence from diagnostic classification. [16–18]

Make the population comparison specific

Ask for separate counts for age band, language use, education, dialect/region and clinical status. “Hispanic,” “Spanish-speaking” and “bilingual” are not interchangeable sampling variables.

Illustrative interpretation: A child knows “bridge” in English and “puente” in Spanish. A prescribed conceptual procedure may credit the concept once. That does not mean the child can explain a bridge, understand a story about it, or use either label across contexts. Assess the actual referral demand.

Structured samples and learning measures

Standardized procedures can be useful even when they do not produce national norm-referenced scores.

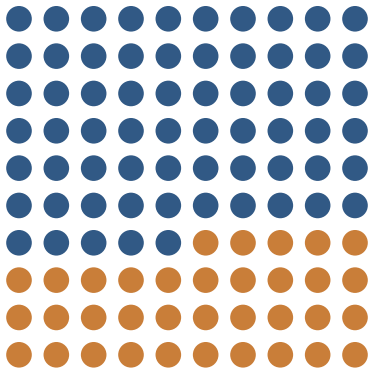
Method	Sample and best use	Do not infer
SALT bilingual retell [19]	TX/CA K–3; overall 5;0–9;9. Reference counts 2,070 / 1,667 / 930 by story group. Match elicitation and transcription.	Counts are story samples, not unique children. New stories do not inherit norms.
MAIN Spanish/English [21]	Originally about 3–10. Structured narrative production/comprehension. Linked Spanish version is European Spanish.	Development pilots are not US bilingual norms. Story/mode choices matter.
SLAM [20]	Select a developmentally and linguistically suitable set for narrative, inference and explanation.	Informal rubric totals are not standardized diagnostic scores.
Word-learning DA [22, 24]	Named study: N=28, ages 4–5; predominantly Spanish-speaking bilinguals. Examine learning under support.	A new task does not inherit study accuracy.
Narrative DA [23, 24]	Named English study: N=54 Spanish-English children; learning and modifiability measures.	There is no single diagnostic accuracy pair for all DA.
Original criterion probe [1]	Define a skill, opportunity, prompt, response and score; monitor change.	An 80% goal criterion is not a diagnosis or an age norm.

Name the evidence correctly

Norm referenced compares performance with a defined reference population. Criterion referenced asks whether a specified skill or criterion is met. Dynamic assessment studies learning under mediation. Informal describes the procedure's status, not whether it is automatically better or worse.

What specificity changes in practice

This illustration explains a metric. It does not predict the result for the next child.



65 test negative

Correct negatives in this illustration

35 test positive

False positives in this illustration

Illustration: among 100 children without the target disorder, 65% specificity would yield 65 correct negatives and 35 false positives. The CELF-4 Spanish study reported 65% specificity in its particular low-income, English-only-school sample. [10]

The denominator matters

This is 35 out of 100 children who do not have the disorder. It is not a claim that 35% of all children tested will be misidentified. Predictive value also depends on how common the disorder is in the group being assessed.

Finding	Reasonable next action
Low Spanish score after years of English-only schooling	Review Spanish exposure and opportunities; obtain relevant samples and family information.
Higher English score and low Spanish score	Explain the uneven profile; evaluate the referred functional concern across contexts.
Difficulty across languages plus concerns in familiar contexts	Integrate language-specific evidence, learning response and functional impact; examine other explanations.

Report language: “The score is one part of the evaluation. The study most relevant to this child’s schooling profile showed reduced specificity. I therefore interpreted it with family report, samples in the child’s languages, and response to support.” Adapt this wording to the evidence actually collected.

Do not use a low score alone to diagnose, or a strong score on one narrow task alone to dismiss a persistent functional concern. [1]

Three referral profiles and a defensible plan

Fictional cases demonstrate decision-making, not preset diagnostic rules.

CASE 1

A preschooler with distributed vocabulary

Age 4;8; Spanish at home, English in preschool; family and teacher both want clearer messages.

Review BESA and PLS-5 Spanish age/profile rules. Pair an appropriate measure with caregiver description, play or narrative samples, hearing/access information and functional observations.

Clinical move: Choose a score only after documenting fit. A bilingual label does not choose the test for you. [1, 2, 5]

CASE 2

An eight-year-old with weak explanations

Spanish with family; English-only schooling since kindergarten. The teacher reports difficulty explaining causes.

BESA's listed age range does not fit. Examine a matched Spanish/English narrative or explanation procedure and relevant language-specific probes. Consider CELF-4 Spanish or English CELF-5 only with explicit norm-fit limits.

Clinical move: Do not replace missing older bilingual norms with a younger test's norms or with CELF-5 Spain. [1, 9, 11]

CASE 3

A child whose speech differs across listeners

Age 6; Caribbean Spanish-speaking family; listeners unfamiliar with the child report reduced intelligibility.

Gather words and connected speech in both languages, examine dialect patterns, and ask familiar listeners about everyday communication. Consider Spanish speech tools, noting that BAPA's Spanish sample was predominantly Mexican Spanish.

Clinical move: A dialect difference and an unfamiliar listener's reaction are not, by themselves, evidence of a speech disorder. [1, 14, 27]

COPY INTO YOUR WORKFLOW

Your assessment evidence note

Complete this before writing the conclusion. Keep identifying information in your approved clinical record system.

Referral question and daily communication demand

Languages by person and setting; schooling; relevant access and dialect information

Evidence item	Exact version score and procedure	Population fit or limitation

Accuracy source: cutoff, clinical N, typical N, reference standard and uncertainty

What findings converge; what conflicts; what evidence is still needed

Quick manual request

Ask the publisher or consult your authorized manual for the exact diagnostic table, norm-group inclusion rules, subgroup counts and permitted accommodations. Missing public information should become a documented question, not an invented value.

Sources for language comparisons

Public publisher information, manuals and original studies. Last checked September 26, 2026.

- [1] ASHA. Multilingual service delivery in audiology and speech-language pathology. Practice Portal.
<https://www.asha.org/practice-portal/professional-issues/multilingual-service-delivery/>
- [2] Brookes Publishing. Bilingual English Spanish Assessment product information. <https://brookespublishing.com/product/besa/>
- [3] Loveall, S. J., Channell, M. M., Mattie, L. J., & Barkhimer, A. E. (2022). Inclusion of individuals with neurodevelopmental disorders in norm-referenced language assessments. *Frontiers in Psychology*, 13, 929433. Table 1. <https://doi.org/10.3389/fpsyg.2022.929433>
- [4] Peña, E. D., Bedore, L. M., & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders*, 51(2), 192–202.
<https://doi.org/10.1111/1460-6984.12199>
- [5] Pearson. Preschool Language Scales Fifth Edition Spanish. Product information.
<https://www.pearsonassessments.com/store/en/usd/p/100000401.html>
- [6] Pearson. PLS 5 family brochure.
<https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/pls5/pls-5-brochure.pdf>
- [7] Pearson. Preschool Language Scales Fifth Edition. US product information and bilingual scoring FAQ.
<https://www.pearsonassessments.com/store/en/usd/p/100000233.html>
- [8] Pearson. PLS 5 Australian and New Zealand Language Adapted Edition. Publisher description identifies the US standardization sample. <https://www.pearsonclinical.com.au/store/en/p/P100010183.html>
- [9] Pearson. CELF 4 Spanish. US product information and comparison with CELF 5 Spain.
<https://www.pearsonassessments.com/store/en/usd/p/100000436.html>
- [10] Barragan, B., Castilla-Earls, A., Martinez-Nieto, L., Restrepo, M. A., & Gray, S. (2018). Performance of low-income dual language learners attending English-only schools on the Clinical Evaluation of Language Fundamentals Fourth Edition, Spanish. *Language, Speech, and Hearing Services in Schools*, 49, 292–305. https://doi.org/10.1044/2017_LSHSS-17-0013
- [11] Pearson. CELF 5 Spain. Product information and comparison of Spanish editions.
<https://www.pearsonassessments.com/store/en/usd/p/P100035000.html>
- [12] Wiig, E. H., Semel, E., & Secord, W. A. (2013). CELF 5 examiner's manual. Pearson. p. 153, diagnostic combinations and cut scores. Public publisher copy.
<https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/field-research/celf5-examiners-manual.pdf>
- [13] Pearson. CELF 5 technical information infographic. Normative N and publisher accuracy table.
https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/celf-5/8439-A_CELF-5_Infographic_Hr_f.jpg
- [28] Pearson. CELF Preschool 2 Spanish. Product information, age range and standardization summary.
<https://www.pearsonassessments.com/store/en/usd/p/100000167.html>

Sources for speech and complementary methods

References support the named claims; original case examples and worksheets are editorial applications.

- [14] Smarty Ears. Bilingual Articulation and Phonology Assessment. Standardization and accuracy table. <https://www.smartyearsapps.com/bilingual-articulation-and-phonology-assessment/>
- [15] Pearson. GFTA 3 Spanish technical summary. Sonidos-en-palabras accuracy table. <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/gfta-3/GFTA-3-Spanish-Technical-Summary.pdf>
- [16] Academic Therapy Publications. EOWPVT 4 Spanish Bilingual Edition. Product and administration information. <https://assessment.s.academictherapy.com/i/expressive-one-word-picture-vocabulary-test-4th-edition-spanish-bilingual-edition-eowpvt-4-sbe>
- [17] Academic Therapy Publications. ROWPVT 4 Spanish Bilingual Edition. Product information. <https://assessments.academictherapy.com/i/receptive-one-word-picture-vocabulary-test-4th-edition-spanish-bilingual-edition-rowpvt-4-sbe>
- [18] De Anda, S., Cysyk, L. M., Moore, H., Huerta, L., Larson, A. L., & King, M. (2022). Psychometric properties of the English Spanish Vocabulary Inventory in toddlers with and without early language delay. *Journal of Speech, Language, and Hearing Research*, 65(2), 672–691. https://doi.org/10.1044/2021_JSLHR-21-00240
- [19] SALT Software. Bilingual Spanish English Story Retell reference databases. Database counts, participants and elicitation protocol. <https://www.saltsoftware.com/coursefiles/SelfCheck/shared/BilingualSEStoryRetellRDBDoc.pdf>
- [20] LEADERSproject. School Age Language Assessment Measures. Materials index. <https://www.leadersproject.org/en/category/slam>
- [21] Gagarina, N., et al. (2019). MAIN Multilingual Assessment Instrument for Narratives revised. *ZAS Papers in Linguistics*, 63. European Spanish version adapted by M. J. Ezeizabarrena (2020). https://www.leibniz-zas.de/fileadmin/media/Dokumente/ZASPiL/ZASPiL_63/MAIN_Spanish_2020_Revised.pdf
- [22] Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech, and Hearing Services in Schools*, 43, 81–96. [https://doi.org/10.1044/0161-1461\(2011/10-0095\)](https://doi.org/10.1044/0161-1461(2011/10-0095))
- [23] Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *Journal of Speech, Language, and Hearing Research*, 57, 2208–2220. https://doi.org/10.1044/2014_JSLHR-L-13-0151
- [24] Orellana, C. I., Wada, R., & Gillam, R. B. (2019). The use of dynamic assessment for the diagnosis of language disorders in bilingual children: A meta-analysis. *American Journal of Speech-Language Pathology*, 28, 1298–1317. https://doi.org/10.1044/2019_AJSLP-18-0202
- [25] ASHA. Collaborating with interpreters, transliterators, and translators. Practice Portal. <https://www.asha.org/practice-portal/professional-issues/collaborating-with-interpreters/>
- [26] Bilinguistics. Evaluation resources. Resource index for further professional learning. <https://bilinguistics.com/evaluation-resources/>
- [27] Pearson. Goldman Fristoe Test of Articulation Third Edition Spanish. Product information. <https://www.pearsonassessments.com/store/usd/p/100001920.html>

Further professional learning: Bilinguistics evaluation resources [26] and LEADERSproject SLAM materials [20]. Availability of a resource does not, by itself, establish diagnostic validity.

Sources for the expanded accuracy tables

Secondary reports are identified explicitly; their limitations remain beside the clinical tables.

[29] LEADERSproject. (2013). Test review: PLS-5 Spanish. pp. 5–6. Secondary manual review.

https://cpb-us-w2.wpmucdn.com/edblogs.columbia.edu/dist/d/730/files/2015/05/PLS-5-Spanish-2012_Pub_10.29.13-18cdkt8.pdf

[30] Bilingualistics. Common communication measures for diverse students. pp. 42–43 (PDF pp. 8–9). Secondary comparison.

<https://bilingualistics.com/wp-content/uploads/2013/12/Common-Communication-Measures-for-Diverse-Students.pdf>

[31] King, M., Larson, A. L., & Buzhardt, J. (2022; online 2021). Exploring the Classification Accuracy of the Early Communication

Indicator (ECI) With Dual-Language Learners From Latinx Backgrounds. *Assessment for Effective Intervention*, 47(4), 209–219. p. 212: secondary PLS attribution. <https://files.eric.ed.gov/fulltext/EJ1347585.pdf>

[32] LEADERSproject. (2013). Test review: PLS-5 English. pp. 5–6. Secondary US manual review.

<https://bpb-us-w2.wpmucdn.com/edblogs.columbia.edu/dist/e/748/files/2015/05/PLS5-English-finaldraft-2dgptop.pdf>

[33] Castilla-Earls, A., et al. (2020). Beyond Scores: Using Converging Evidence to Determine Speech and Language Services

Eligibility for Dual Language Learners. *AJSLP*, 29, 1116–1132. p. 1126: secondary accuracy reports. <https://iilta.org/wp-content/uploads/2020/12/Beyond-Scores-Using-Converging-Evidence-to-Determine-Speech-and-Language-Services-Eligibility-for-Dual-Language-Learners.pdf>

[34] Lazewnik, R., et al. (2019). Identifiers of Language Impairment for Spanish-English Dual Language Learners. *LSHSS*, 50(1),

126–137. Primary abstract verifies study context. <https://repository.daemen.edu/works/publication-article/8esdw-e6f37>

[35] LEADERSproject. (2014). Test review: CELF-4 Spanish. pp. 7–8. Secondary manual review.

https://cpb-us-w2.wpmucdn.com/edblogs.columbia.edu/dist/d/730/files/2015/05/TestReview_CELF-4-Spanish_1.8.14-1z5jjx9.pdf

Sources for the additional English measures

The catalog retains fuller source links and study notes for maintenance.

[36] Pearson. CELF-5 Spain brochure. p. 6, normative sample.

<https://www.pearsonclinical.es/content/dam/school/global/clinical/clinical-es/whitepapers/Folleto-CELF-5-Spain-WEB.pdf>

[37] Pearson. (2020). CELF Preschool-3 technical summary. Accuracy table and norms.

<https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/celf-preschool-3/celf-p3-technical-summary.pdf>

[38] Nelson, N. W., Plante, E., Helm-Estabrooks, N., & Hotz, G. (2016). TILLS technical manual excerpt. Brookes. pp. 10–11, Tables 2.1–2.2. <https://tillstest.com/wp-content/uploads/2021/04/TILLS-Technical-Chapter-excerpt.pdf>

[39] Brookes Publishing. TILLS frequently asked questions. Language-experience eligibility and use. <https://tillstest.com/faqs/>

[40] Ferrell, K. R. (n.d.). Understanding sensitivity and specificity. WPS. Table 2, p. 5, original CASL-2 analysis.

https://ecom-cdn.wpspublish.com/prod/media/wysiwyg/casl-2_sensitivity_specificity_1.pdf

[41] WPS. (2024, July 30). Oral language assessment webinar handout. Slides 63–66, CASL-2 norms and English proficiency.

<https://pages.wpspublish.com/hubfs/Webinar%20PPT%20Slides/Oral%20Language%20PPT%20handout%20073024.pdf>

[42] Brookes Publishing. (2015). How to get the most out of TILLS: Webinar Q&A. p. 3, normative N.

<https://brookespublishing.com/wp-content/uploads/2021/06/TILLS-How-to-get-the-most-out-of-TILLS-webinar-QA-2015.pdf>

Source practice: missing diagnostic denominators and confidence intervals are not estimated from rounded rates. A norm sample, a development sample and an accuracy-analysis sample are different counts. Figures were checked September 27, 2026; confirm the edition and licensed manual before operational use.