ChatGPT

Interpretación profesional de la CELF-5 (versión española, Pearson)

Descripción general de la CELF-5 (España)

La **CELF-5 española** (Evaluación Clínica de los Fundamentos del Lenguaje, 5ª Edición) es una batería estandarizada diseñada para evaluar las habilidades lingüísticas de niños y adolescentes de 5 años hasta 15 años 11 meses [1] [2]. Está compuesta por **12 subpruebas** que miden distintos componentes del lenguaje (semántica, léxico, morfología, sintaxis, pragmática y memoria verbal) [3] [4]. Estas subpruebas incluyen, por ejemplo:

- **Comprensión de frases (CF)** – Comprensión auditiva de oraciones sencillas.
- **Conceptos lingüísticos (CL)** – Comprensión de conceptos básicos y relaciones (solo para 5-8 años) [5].
- **Morfosintaxis (M)** – Uso de morfología y sintaxis en expresiones orales.
- **Palabras relacionadas (PR)** – Relación y clasificación semántica de palabras.
- **Ejecución de indicaciones (EI)** – Seguimiento de instrucciones verbales complejas (memoria auditiva y comprensión).
- **Elaboración de frases (EF)** – Formulación de oraciones a partir de estímulos (evalúa expresión sintáctica).
- **Repetición de frases (RF)** – Repetición literal de oraciones (memoria verbal y sintaxis).
- **Comprensión oral de textos (COT)** – Comprensión de narraciones orales (solo aplicable en niños mayores).
- **Definición de palabras (DP)** – Explicación del significado de palabras (vocabulario expresivo; solo >8 años) [5].
- **Puzle de palabras (PP)** – Organización de palabras para formar oraciones lógicas (gramática/semántica; solo >8 años) [5].
- **Relaciones semánticas (RS)** – Identificación de relaciones lógicas entre palabras/frases (solo >8 años) [5].
- **Perfil de habilidades pragmáticas (PHP)** – Cuestionario de observación de la comunicación pragmática (rellenado por el profesional mediante la observación y colaboración de padres/maestros) [6].

Las primeras 11 subpruebas se administran directamente al alumno, mientras que la última (PHP) la completa el evaluador a partir de la observación y de información de padres o profesores [6]. La batería se divide en dos rangos de edad (5-8 años y 9-15 años); algunas subpruebas se aplican solo en un rango (por ejemplo, CL solo en menores, y DP/RS/PP solo en mayores) [7]. Además de las subpruebas, la CELF-5 ofrece **recursos complementarios**: la **Verificación de habilidades pragmáticas (VHP)**, una lista de chequeo que permite evaluar de forma criterial las conductas pragmáticas observadas, para apoyar el diagnóstico de trastorno pragmático del lenguaje [8]; y un **Cuestionario de competencia lingüística (CCL)** para padres/educadores, que aporta una visión del desempeño comunicativo del niño en entornos cotidianos (escolar y familiar) [9]. Estos recursos contextuales ayudan a integrar la información de la prueba con la funcionalidad del lenguaje en la vida diaria.

Puntuaciones estándar y baremos en la CELF-5

Estandarización: La CELF-5 española cuenta con baremos tipificados en población infantil y juvenil de España, lo que permite comparar el rendimiento de un alumno con el de sus pares por edad. Las puntuaciones de cada subprueba se convierten en **puntuaciones escalares estandarizadas** con media 10 y desviación típica (DT) 3 [10] [11]. Por su parte, las puntuaciones combinadas de ciertas subpruebas proporcionan **índices compuestos** (como Índice de Lenguaje Receptivo, Expresivo, etc.) expresados como puntuaciones típicas con media 100 y DT 15 [10]. También se ofrecen **percentiles** asociados y, en casos específicos, equivalencias de edad y **valores de desarrollo** (escala Rasch con media 500, DT 25) para seguimiento del progreso [10]. En resumen, los tipos de puntuaciones disponibles incluyen: puntuaciones directas (brutas), escalares, compuestas, percentiles, rangos de confianza, edades equivalentes (excepto en algunas pruebas como comprensión de textos) y puntuaciones de desarrollo [12] [13].

Puntuaciones escalares (subpruebas): Cada subprueba produce una puntuación escalar de 1 a 19 (media = 10, DT = 3) basada en la conversión de la puntuación directa según la edad del sujeto [14] [15]. Una puntuación escalar de **10** refleja el rendimiento medio esperado para la edad. Las puntuaciones escalares de **7** o **13** corresponden aproximadamente a 1 desviación típica por debajo o por encima de la media, respectivamente [16]. De hecho, alrededor de dos tercios (~66%) de los niños con desarrollo lingüístico normal obtienen puntuaciones escalares entre 7 y 13 en las subpruebas [16]. Esto significa que rangos de puntuación dentro de ese intervalo pueden considerarse dentro de la variabilidad normal esperada.

Índices compuestos: La CELF-5 define varios índices compuestos obtenidos sumando las puntuaciones escalares de subpruebas específicas y convirtiendo el total en una escala típica (media = 100, DT = 15) [17] [18]. Los **5 índices principales** son: la **Puntuación Principal de Lenguaje (PPL)**, que resume la aptitud lingüística general; el **Índice de Lenguaje Receptivo (ILR)**, que refleja la comprensión auditiva; el **Índice de Lenguaje Expresivo (ILE)**, que refleja la capacidad de expresión oral; el **Índice de Contenido Lingüístico (ICL)**, centrado en la dimensión léxico-semántica; y el **Índice de Estructura Lingüística (IEL)**, centrado en la dimensión morfosintáctica (en el grupo de mayor edad, el IEL pasa a llamarse Índice de Memoria Lingüística, asociado a tareas de memoria verbal y sintaxis) [19]. Cada índice compuesto se calcula a partir de un conjunto fijo de subpruebas: por ejemplo, la PPL se obtiene combinando cuatro subpruebas troncales (p.ej., Comprensión de frases, Morfosintaxis, Elaboración de frases y Repetición de frases) [20], mientras que el ILR integra típicamente pruebas receptivas (ej. Comprensión de frases, Palabras relacionadas, Ejecución de indicaciones) [21], etc. Todos los índices se escalan a media 100 (rendimiento promedio esperado para la edad) y DT 15 [22] [23], facilitando así la interpretación de qué tan por encima o por debajo del promedio se sitúa el desempeño lingüístico global o en dominios específicos del niño.

Interpretación de las puntuaciones escalares (subpruebas)

Las puntuaciones escalares de cada subprueba permiten valorar de manera **descriptiva** el nivel de habilidad específica del niño en esa área del lenguaje, comparado con sus pares. Para interpretar correctamente estas puntuaciones, se emplean **rangos descriptivos** asociados a la distancia en desviaciones típicas respecto a la media. En la versión española de la CELF-5 se utiliza la siguiente clasificación para las puntuaciones escalares individuales [24]:

- **Por encima de la media:** Puntuación escalar de **13 o superior** (≥ +1 DT). Indica un rendimiento **elevado** en la habilidad evaluada, superior al esperado para la edad (aprox. percentil 84 o mayor) [25] [24]. Supone una fortaleza relativa del niño en esa área del lenguaje.

- **En la media:** Puntuación escalar de **8 a 12** (entre +1 y –1 DT). Indica un rendimiento **promedio** o típico, dentro de lo esperado para la edad (entre los percentiles ~18 y 83, abarcando la mayor parte de los niños normativos) [26] [24]. Implica que la habilidad evaluada está desarrollada adecuadamente sin indicios de rezago significativo.
- **En el límite:** Puntuación escalar de **7** (≈ –1 DT). Representa un rendimiento **borderline**, justo en el límite inferior de lo normal [24]. Equivale aproximadamente al percentil 16 (1 desviación bajo la media); suele interpretarse con precaución, ya que puede indicar una leve dificultad en esa área, aunque podría aún caer dentro de la variabilidad normal. Es un nivel a vigilar, pues está en el borde entre lo típico y una posible deficiencia leve.
- **Por debajo de la media:** Puntuación escalar de **6 o inferior** (< –1 DT). Indica un rendimiento **bajo** en la habilidad específica, significativamente por debajo del promedio de la edad [24]. Por ejemplo, una puntuación de 4 correspondería a –2 DT (percentil ~2) y una de 1 a –3 DT (percentil ~0.1) [25]. Estos valores implican **dificultades notorias** en el área evaluada por la subprueba. Cuanto más se aleje la puntuación por debajo de 7, más grave es el déficit: 6 sería leve-bajo, mientras 1-3 denotan un rendimiento **muy bajo**. En informes, suele resumirse simplemente como "por debajo de la media" para cualquier puntuación ≤6, aunque el evaluador debe apreciar la gradación de la severidad dentro de este rango (p.ej., 5-6 = déficit leve; 1-4 = déficit moderado a severo).

Estas categorías ayudan a describir de forma clara el desempeño del niño en cada habilidad lingüística específica. Por ejemplo: *"En la subprueba X (puntuación escalar = 5), el niño obtuvo un rendimiento por debajo de la media para su edad, lo que sugiere una debilidad significativa en... (esa habilidad)"*. En cambio, una puntuación de 11 se describiría como *"en la media"* (desempeño adecuado). Esto facilita la comunicación de resultados a padres y docentes en términos comprensibles [27] [28]. Siempre es importante recordar que una puntuación individual ligeramente baja en una subprueba debe interpretarse dentro del patrón general de resultados: un área débil aislada podría compensarse con otras habilidades, mientras que múltiples subpruebas por debajo de la media apuntan a un perfil de dificultades más amplio.

Interpretación de los índices compuestos (puntuaciones compuestas)

Los **índices compuestos** brindan una visión global de diferentes dominios del lenguaje, integrando varias subpruebas, por lo que suelen ser aún más informativos para diagnóstico que las puntuaciones de una sola prueba [29]. Al interpretar los índices (que son puntuaciones típicas con media 100), se emplean también categorías descriptivas basadas en la desviación estándar respecto a la media [30] [26]. La clasificación utilizada en la CELF-5 española es la siguiente:

- **Por encima de la media:** Índice estándar **≥ 115** (≥ +1 DT por encima de la media). Refleja un rendimiento **superior** al promedio de la edad en el área evaluada por ese índice compuesto (percentil ~84 o superior) [26]. Indica que el niño manifiesta habilidades lingüísticas notablemente mejores que las típicas para su grupo de edad en ese dominio (ej. comprensión o expresión sobresaliente).
- **En la media:** Índice entre **86 y 114** (dentro de ±1 DT). Corresponde a un rendimiento **promedio** o **normal** en la habilidad global medida (percentil aproximadamente 18 a 83) [26]. Un niño en este rango no presenta indicios de rezago lingüístico significativo en ese ámbito; su desempeño se considera típico para la edad.
- **En el límite:** Índice entre **78 y 85** (entre –1 y –1.5 DT por debajo de la media). Indica un rendimiento **límite/borderline**, ligeramente por debajo de lo esperable, que podría sugerir dificultades leves (percentil ~7 a 17) [31]. Los niños en este rango pueden tener algunas *señales*

de dificultad lingüística incipiente, pero no necesariamente cumplen criterio de trastorno; requieren observación adicional. Este rango limítrofe coincide aproximadamente con hasta –1.5 DT, umbral a partir del cual suele considerarse clínicamente relevante la desviación.

- **Por debajo de la media:** Índice entre **71 y 77** (entre –1.5 y –2 DT). Representa un rendimiento **bajo** (percentil ~3 a 6) claramente inferior al promedio [31]. Implica **dificultades significativas** en el dominio lingüístico evaluado; típicamente se consideraría que el niño tiene un trastorno moderado del lenguaje en esa área.
- **Muy por debajo de la media:** Índice $\leq$ **70** (–2 DT o más por debajo de la media). Refleja un rendimiento **muy bajo** (extremadamente por debajo del promedio, percentil ~2 o menos) [32]. Este nivel de puntuación indica **déficits severos** en el lenguaje. Niños con puntuaciones compuestas en este rango suelen presentar trastornos del lenguaje graves, que requieren intervención inmediata.

Estas categorías ofrecen descriptores cualitativos útiles para comunicar la gravedad del trastorno. Por ejemplo, un **Índice de lenguaje expresivo (ILE)** de 80 caería *"en el límite"* (aprox. –1.3 DT, percentil ~9) sugiriendo un posible problema leve en la expresión, mientras que un **Índice de lenguaje receptivo (ILR)** de 70 sería *"muy por debajo de la media"* (–2 DT, percentil ~2) indicando una **afectación severa** de la comprensión [33]. En informes profesionales, a menudo se asocian estos rangos con etiquetas de severidad clínica: p.ej., **leve** (déficit alrededor de 1–1.5 DT bajo la media), **moderado** ($\approx$1.5–2 DT bajo la media) o **severo** (>2 DT por debajo) [34] [35].

Es importante destacar que, según los **estudios de validación españoles**, un **índice compuesto de 85** (justo 1 DT por debajo de la media) resultó ser un **punto de corte clínico útil**: en una muestra española de niños con trastorno del lenguaje versus grupo control, usar **85** como valor umbral permitió identificar correctamente al **100%** de los niños con trastorno y sin falso positivos en el grupo control (sensibilidad = 1.00, especificidad = 1.00) [36]. En otras palabras, **un índice $\leq$85** en la Puntuación Principal de Lenguaje (PPL) o en los índices Receptivo/Expresivo es fuertemente indicativo de trastorno lingüístico, según la evidencia normativa española [36]. Por ello, la guía técnica sugiere que si la PPL, ILR o ILE del niño es **85 o menor**, se **amplíe la evaluación** aplicando todas las subpruebas adicionales para detallar el perfil de debilidades específicas [37]. Este procedimiento ayuda a confirmar el diagnóstico y precisar en qué componentes del lenguaje el niño muestra más dificultades.

Sin embargo, resultados en el rango limítrofe (por ejemplo índices entre 78–85) deben interpretarse con cautela. Un niño con puntuaciones compuestas borderline puede no cumplir criterios estrictos de trastorno, pero sí presentar vulnerabilidades que podrían impactar su desempeño académico o comunicativo. En estos casos, conviene valorar otras evidencias: observaciones clínicas, informes escolares, progreso con el tiempo, etc., antes de decidir si requiere intervención.

Significado clínico de los rangos de puntuación

Los rangos de puntuación anteriores tienen implicaciones directas en la **conclusión diagnóstica** y en la planificación de apoyos:

- Un perfil de **puntuaciones "en la media"** en todos los índices sugiere que el niño **no presenta un trastorno del lenguaje** evidente. Su competencia lingüística general está dentro de lo esperable para su edad, aunque podría haber pequeñas variaciones individuales. En ámbito escolar, esto implicaría que no requiere intervención especial de lenguaje, salvo seguimiento rutinario.
- **Puntuaciones "por encima de la media"** evidencian áreas de fortaleza. Por ejemplo, un niño puede tener un índice de contenido lingüístico muy alto, indicando vocabulario rico y buen razonamiento semántico. Estas fortalezas pueden compensar ciertas debilidades y sugieren un

potencial elevado; en contexto educativo puede significar que el alumno destaca en tareas verbales complejas.

- **Resultados "en el límite"** (borderline) señalan que el niño está cerca de la frontera de la normalidad y la patología. Clínicamente, esto podría interpretarse como un posible **trastorno leve** o simplemente una variación normal baja. En la práctica, se deben considerar factores como: ¿El niño tiene dificultades funcionales evidentes en el aula o en casa? ¿Existe historial de problemas de lenguaje en desarrollo temprano? Un desempeño limítrofe requiere quizás **monitoreo continuo** y, si es posible, prevención (ej. estrategias de refuerzo en el aula) para evitar que se agrave. Muchos manuales consideran que puntuaciones alrededor de –1,25 DT (entre 80–85) son criterios de *"risk zone"* o riesgo de TEL, dependiendo de otras evidencias [38] [39].
- **Resultados "por debajo de la media" o "muy por debajo"** implican **déficit significativo**. Un niño con índices compuestos en 70–77 (bajo) o ≤70 (muy bajo) casi con certeza cumple criterios diagnósticos de un **Trastorno del Lenguaje** (sea Trastorno del Lenguaje de desarrollo – antes llamado TEL, o asociado a otras condiciones) [38] [40]. La **severidad** del trastorno se gradúa precisamente por la distancia a la media: por ejemplo, una PPL ~75 sugiere un trastorno moderado, mientras ~60 o menos indicaría un trastorno severo que afecta profundamente la comunicación. En el contexto escolar, estos alumnos suelen requerir adaptaciones curriculares, apoyo de especialistas (logopedia) e intervenciones focalizadas. En contexto clínico, ameritan un diagnóstico formal y tratamiento terapéutico.

En síntesis, **puntuaciones compuestas ≤ –1 DT (≤85)** deben alertar al evaluador de una posible patología del lenguaje. Cuanto mayor es la desviación negativa, mayor es la **probabilidad y gravedad** del trastorno. Estas interpretaciones están alineadas con los criterios diagnósticos habituales: muchos profesionales toman –1,5 DT (~puntuación 78) como punto de corte para diagnosticar Trastorno del Lenguaje (especialmente en investigación), aunque en la práctica se combina con juicio clínico [38]. La **CELF-5 española**, gracias a su baremación local, muestra una excelente capacidad diagnóstica para detectar niños con TEL usando dichos cortes [41]. Se ha señalado que es una de las herramientas **más sensibles y específicas** para diagnóstico de trastorno específico del lenguaje (TEL/DLD) en español [40], siempre que se apliquen adecuadamente sus criterios.

Integración de los resultados para conclusiones diagnósticas

Una interpretación profesional de la CELF-5 va más allá de asignar etiquetas a puntuaciones: requiere integrar todos los hallazgos para obtener un **perfil lingüístico coherente** y llegar a conclusiones diagnósticas sólidas tanto en el ámbito escolar como clínico. A continuación, se detallan consideraciones para dicha integración:

1. Análisis del perfil de índices compuestos: Es importante comparar entre sí los distintos índices del niño, buscando discrepancias significativas que aporten información diagnóstica. La CELF-5 provee valores críticos para determinar cuándo una diferencia entre dos índices es estadísticamente significativa/unusual en la población. Por ejemplo, una diferencia de ~13 puntos o más entre el Índice de Lenguaje Receptivo y el Expresivo es poco común en la muestra normativa [42], por lo que si un niño muestra ILR mucho menor que ILE, sugiere un perfil inusual donde la **comprensión auditiva está notablemente más afectada que la expresión** (posible indicio de un trastorno receptivo). Inversamente, un ILE bastante inferior al ILR indicaría dificultades expresivas marcadas en contraste con una comprensión relativamente conservada, perfil típico de algunos trastornos expresivos. De igual modo, diferencias significativas entre el **Índice de Contenido (ICL)** y el **Índice de Estructura (IEL/IML)** pueden revelar la naturaleza del trastorno: p. ej., un ICL mucho más bajo que el IEL apuntaría a un déficit léxico-semántico (vocabulario pobre, dificultades de significado) con habilidades morfosintácticas relativamente mejores; mientras que un IEL menor que ICL sugeriría mayores dificultades gramaticales/

morfosintácticas en comparación con el conocimiento de vocabulario [43]. La **magnitud de la diferencia** y su frecuencia en la población normativa determinan si es clínicamente relevante: la CELF-5 proporciona tablas de *diferencias críticas* (con intervalos de confianza) para guiar esto [44] [45]. El evaluador debe comentar cualquier brecha significativa entre índices, ya que orienta las **necesidades específicas** de intervención (ej.: "discrepancia significativa comprensión-expresión, con comprensión muy reducida: el niño puede entender menos de lo que logra verbalizar, afectando su seguimiento de instrucciones en clase").

2. Patrones en las subpruebas: Dentro de cada índice o área, conviene identificar **subpruebas particularmente bajas o altas**. Las subpruebas más deficitarias señalan **puntos débiles específicos**. Por ejemplo, en un niño con PPL baja, determinar *qué* subpruebas contribuyeron a esa baja puntuación es crucial: ¿Fueron principalmente las pruebas de memoria (p. ej. Repetición de frases, Ejecución de indicaciones) las que hundieron el puntaje? ¿O fueron las de morfología/sintaxis (Morfosintaxis, Elaboración de frases) las más bajas? Esto diferencia un posible trastorno de memoria verbal secuencial de un trastorno morfosintáctico. Del mismo modo, comparar desempeño en **Comprensión de textos orales** vs **Comprensión de frases** puede indicar si el niño tiene más problemas en comprender información extensa o inferencial que oraciones simples. Un análisis cualitativo de los errores del niño en las subpruebas (por ejemplo, tipos de estructuras gramaticales que no manejó, clases de palabras que no comprendió, etc.) añade valor diagnóstico, aunque ese nivel de detalle va más allá de la puntuación numérica. En contextos escolares, vincular estas debilidades con las exigencias del currículo es útil: por ejemplo, un niño con baja *Ejecución de indicaciones* tendrá problemas para seguir instrucciones multilaterales en clase, y uno con baja *Elaboración de frases* puede tener dificultades para construir oraciones correctamente en la expresión escrita.

3. Evidencias pragmáticas y funcionales: La CELF-5 incorpora un Perfil de Habilidades Pragmáticas (PHP) y la lista de Verificación de Habilidades Pragmáticas (VHP) que evalúan la **comunicación social** del niño [8]. A diferencia de las otras pruebas, el resultado pragmático no se da en puntuaciones típicas sino de forma criterial (se compara con un punto de corte) [46] [47]. Si el niño obtiene una puntuación por debajo del criterio en la VHP (por ejemplo, dificultades marcadas en uso del lenguaje en contexto, responder de forma no literal, tomar turnos, etc.), esto sugiere la presencia de un **trastorno pragmático de la comunicación** (como un Trastorno de la Comunicación Social) o rasgos pragmáticos de un TEL. Es fundamental integrar estos hallazgos con los índices formales: puede haber niños con buen desempeño estructural (gramatical) pero con comunicación inadecuada en lo social. En informes, se debe señalar si el **perfil pragmático** está alterado, ya que requiere intervenciones diferenciadas (ej. entrenamiento en habilidades sociales, no solo terapia de lenguaje estructural). Asimismo, el **Cuestionario de competencia lingüística** aportará información del entorno natural del niño – por ejemplo, los padres/maestros pueden reportar que *"no entiende en clase cuentos largos"* o *"le cuesta encontrar palabras para expresarse"*. Estas observaciones contextualizan las puntuaciones y sirven de evidencia **ecológica** de la afectación: refuerzan la fiabilidad diagnóstica al demostrar que las dificultades medidas por la CELF-5 tienen un impacto real en la vida diaria y el aprendizaje del niño [9] [48]. Si, por el contrario, las puntuaciones son bajas pero padres y profesores no perciben problemas funcionales, el evaluador debe considerar posibles **factores atenuantes** o la necesidad de más evaluación (¿fue el niño cooperativo en la prueba? ¿existe alguna condición extralingüística, como atención, que haya influido?).

4. Uso de intervalos de confianza y error de medida: La interpretación profesional siempre tiene en cuenta la **incertidumbre de la medida**. La CELF-5 proporciona errores típicos de medida (ETM) e intervalos de confianza para cada puntuación compuesta y escalar según la edad [49] [50]. Es recomendable reportar, por ejemplo: *"Índice de lenguaje expresivo = 80, IC 90% [74–86]"*. Esto significa que, dado el error estándar, la puntuación "verdadera" del niño probablemente esté en ese rango. Si un índice cae justo en zona limítrofe (por ej. 85) con un intervalo que se extiende hacia 90, conviene ser

prudente antes de etiquetar un trastorno – podría ser que su habilidad real estuviera cerca del promedio. Del mismo modo, intervalos pueden clarificar si una diferencia entre índices es estadísticamente significativa o no. El profesional debe integrar estos datos de fiabilidad para evitar sobreinterpretar pequeñas desviaciones. En general, **decisiones clínicas** (diagnóstico sí/no) no deben basarse en diferencias de unos pocos puntos, sino en tendencias claras considerando los intervalos de confianza [49]. Por ejemplo, si PPL = 82 IC95% [77–87], está casi seguramente por debajo de 85 de forma consistente, apoyando la identificación de un trastorno; pero si PPL = 85 IC95% [79–91], podría no diferenciarse significativamente de la media baja.

5. Consideración de otras pruebas y fuentes: Aunque esta investigación se centra en la CELF-5, en la práctica diagnóstica se suelen combinar **varias pruebas de lenguaje** (p. ej., pruebas de vocabulario como el *PPVT* o de procesamiento fonológico, etc.). Un sistema GPT o profesional que integre datos de múltiples instrumentos debe buscar **convergencia de evidencia**: por ejemplo, si la CELF-5 indica puntuaciones bajas en morfosintaxis y además en otra prueba de lenguaje espontáneo el niño comete muchos errores gramaticales, la conclusión de "trastorno morfosintáctico" se robustece. Si hubiera discrepancias (CELF-5 baja pero otra prueba normal), habría que investigar las razones (diferencias en lo que mide cada test, condiciones de aplicación, etc.). Asimismo, siempre se deben descartar factores auditivos, cognitivos o emocionales que puedan afectar el lenguaje antes de concluir un diagnóstico de trastorno primario del lenguaje. Esto implica integrar información de audiometrías, evaluaciones cognitivas (p. ej. escalas de inteligencia), historiales médicos y contexto socioeducativo. Solo con este **enfoque holístico** se llegará a conclusiones diagnósticas **fiables**.

6. Prudencia y juicio clínico: Finalmente, toda la interpretación debe filtrarse por el **juicio profesional**. Los manuales de la CELF-5 enfatizan que las puntuaciones son una herramienta más y no sustituyen la interpretación experta [27]. Los criterios específicos para diagnosticar un trastorno del lenguaje pueden variar según distintas guías o comunidades autónomas, pero en todos los casos el profesional necesita correlacionar los números con la realidad del niño. Se recomienda comunicar resultados en términos comprensibles, explicando qué significa "por debajo de la media" en palabras funcionales (*"le costará entender instrucciones complejas, por ejemplo..."*). También se aconseja explicitar la **confianza** en el diagnóstico dado el conjunto de datos: por ejemplo, *"los resultados de CELF-5, junto con las observaciones en aula, indican con claridad la presencia de un trastorno del lenguaje de grado moderado en el área expresiva"*. En cambio, si los hallazgos son ambiguos, es válido concluir *"de momento no se confirma un trastorno, pero se observará evolución"*. El uso de los baremos españoles de la CELF-5 aporta objetividad, pero la **decisión final** recae en la síntesis que haga el especialista de **todas las fuentes de información disponibles** [27] [51].

7. Limitaciones de la CELF-5 y necesidad de información complementaria: Es relevante notar que, si bien la CELF-5 es muy completa evaluando niveles léxicos y oracionales, **no aborda en profundidad el nivel discursivo** (narrativo o conversacional extendido) [52]. Por diseño, sus pruebas llegan hasta comprensión y formulación de oraciones, pero no evalúan cómo el niño estructura un relato espontáneo, cómo explica eventos o argumenta ideas en textos más largos [53]. Por tanto, **no se debe emitir un juicio global sobre el lenguaje únicamente con los resultados de la CELF-5** [54]. Un niño podría obtener puntuaciones dentro de la media en oraciones aisladas pero tener dificultades en el lenguaje funcional de la vida real (ej.: para mantener una conversación coherente, relatar una historia con secuencia lógica, etc.). Por ello, en contexto clínico-escolar, se recomienda complementar la CELF-5 con análisis de muestras de lenguaje espontáneo, evaluación de narrativa, observaciones estructuradas en el aula, etc. De hecho, el manual técnico sugiere que la CELF-5 **no sustituye una evaluación pragmática/discursiva más amplia** y que esas dimensiones deben ser examinadas con otros instrumentos o métodos de observación prolongada [55] [54]. Integrar esta información adicional garantizará conclusiones diagnósticas más confiables y completas.

En resumen, la interpretación profesional de la CELF-5 (edición española) requiere combinar los datos normativos (puntuaciones estándar, percentiles y rangos descriptivos) con el análisis cualitativo del perfil de resultados y la información contextual del niño. Siguiendo las directrices de la documentación científica y técnica disponible, el evaluador puede **cuantificar** la magnitud de las dificultades (usando medias, DT, percentiles) y **cualificar** su significado clínico (leve, moderado, severo; área específica afectada), para finalmente **integrar** todo ello en un diagnóstico o informe comprensible. Gracias a la sólida estandarización española de la CELF-5, disponemos de referencias confiables para distinguir rendimiento normal de patológico y para describir con precisión cada subcomponente lingüístico [56] [11]. Utilizada de forma rigurosa, apoyada exclusivamente en fuentes profesionales y científicas, la CELF-5 española se convierte en una herramienta de gran valor para alcanzar **conclusiones diagnósticas fiables** en el ámbito escolar y clínico sobre los trastornos del lenguaje y la comunicación.

Fuentes: La información anterior se ha extraído y sintetizado de manuales técnicos oficiales, evaluaciones psicométricas publicadas y literatura académica sobre la adaptación española de la CELF-5 [56] [4] [24] [26] [53], garantizando que las pautas de interpretación estén respaldadas por evidencia científica y ajustadas a la población española. Todas las referencias citadas corresponden específicamente a la versión española de la CELF-5 y sus estudios de validación, tal como se requería. Esto asegura que las **instrucciones profesionales** aquí recopiladas sean precisas, actualizadas y directamente aplicables en entornos clínicos y educativos de habla hispana.

---

[1] [4] [5] [6] [7] [8] [9] [10] [19] [38] [39] [40] [41] [48] [51] [52] [53] [54] [55] [56] Microsoft Word - Pearson_CELF5_CV_ed_epw
https://www.cop.es/uploads/PDF/2019/CELF-5.pdf

[2] [3] CELF-5 Test de Evaluación Clínica del Lenguaje | Pearson Clinical - Pearson Clinical Assessment España
https://www.pearsonclinical.es/celf-5-evaluacion-clinica-de-los-fundamentos-del-lenguaje-5?srsltid=AfmBOormfO8Mt5yNuQJwmFEApa9hvp-bMsLHvfI49jFPkakN4hZbk3no

[11] [16] [24] [25] [26] [27] [28] [30] [31] [32] [33] [34] [35] [36] [37] pearsonclinical.es
https://www.pearsonclinical.es/mwdownloads/download/link/id/1376?srsltid=AfmBOopjArJkjKKjDLty6-X9vxYFWskhy7y04RAE7gOu8uUUTtnQU_c7

[12] [13] [14] [15] [17] [18] [42] [43] [44] [45] [46] [47] [49] [50] CELF-5 Manual Técnico | PDF | Evaluación | Lingüística
https://es.scribd.com/document/725838821/CELF-5-Manual-Tecnico

[20] [21] [22] [23] pearsonclinical.es
https://www.pearsonclinical.es/mwdownloads/download/link/id/922/?srsltid=AfmBOopFpaCIVtOrtG9QBjDKhLs7dbbQ4BafXjacn01rdHQ5CZkJD1kj

[29] Interpretación de la puntuación CELF-5 para un niño de 11 años : r/slp
https://www.reddit.com/r/slp/comments/1ej2oa6/celf5_scoring_interpretation_for_11yo/?tl=es-es