Revista Iberoamericana de Investigación en Educación
ISSN: e2789-3499 ISSN: 2790-2021

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

https://riied.org

Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado

Effect of pedagogical feedback on English grammar learning in eighth grade

Efeito do feedback pedagógico na aprendizagem de gramática inglesa no oitavo ano

Artículo de investigación
Recibido: 19/10/2026 **Revisado:** 31/10/2026 **Aceptado:** 07/11/2026

Zoila Josefina Torres García
Institución Educativa Municipal, Ciudad de Pasto; Nariño, Colombia
https://orcid.org/0000-0002-1427-1114
torresoila904@gmail.com

Resumen

Objetivo: Analizar los efectos de la especificidad (específica o genérica), la latencia (corta o demorada) y la valencia (neutra o apetitiva) de la retroalimentación pedagógica sobre la ganancia en el aprendizaje de adjetivos, comparativos y superlativos en inglés en estudiantes de octavo grado. **Métodos:** Estudio cuantitativo, cuasiexperimental, con diseño factorial 2×2×2, medición preprueba y postprueba y ocho grupos intactos de una institución pública de San Juan de Pasto (Colombia), cada uno con una condición. Se analizaron los datos de 269 de los 279 estudiantes matriculados. El aprendizaje se midió con una prueba de 25 ítems elaborada para el estudio (α = .77 en la preprueba y .76 en la postprueba). Los datos proceden de la tesis doctoral de la autora y su contraste global entre condiciones se publicó en Torres-García (2019); este artículo incorpora el análisis factorial 2×2×2, el informe de n y medias por condición y una descripción ampliada del método. **Resultados:** Un ANOVA de un factor mostró diferencias entre las ocho condiciones, $F(7, 261) = 2.54$, $p = .015$, $\eta^2 = .06$. El ANOVA factorial 2×2×2 mostró un efecto principal de la especificidad, $F(1, 261) = 9.33$, $p = .002$, η^2 parcial = .035, a favor de la retroalimentación específica; los efectos principales de la latencia, $F(1, 261) = 0.97$, $p = .325$, y de la valencia, $F(1, 261) = 2.89$, $p = .090$, así como todas las interacciones ($p \geq .186$), no fueron significativos. La ganancia media fue mayor en la condición específica, apetitiva y corta (4.8 puntos) y menor en la genérica, apetitiva y demorada (2.0); la diferencia fue significativa en la prueba de Tukey ($p = .041$). **Conclusiones:** Los datos son compatibles con que la retroalimentación específica se asocie con mayores ganancias y no muestran ventaja de los estímulos apetitivos. Con un solo grupo por condición, las diferencias no pueden atribuirse a la retroalimentación con independencia del grupo.

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

https://riied.org

Palabras clave: Retroalimentación pedagógica; Inglés como lengua extranjera; Latencia; Diseño cuasiexperimental; Educación secundaria.

Abstract

Objective: To analyze the effects of the specificity (specific or generic), latency (short or delayed), and valence (neutral or appetitive) of pedagogical feedback on gains in learning adjectives, comparatives, and superlatives in English among eighth-grade students. **Methods:** Quantitative, quasi-experimental study with a 2×2×2 factorial design, pretest-posttest measurement, and eight intact groups from a public school in San Juan de Pasto (Colombia), each assigned one condition. Data from 269 of the 279 enrolled students were analyzed. Learning was measured with a 25-item test developed for the study (α = .77 at pretest and .76 at posttest). The data come from the author's doctoral thesis, and their overall comparison across conditions was published in Torres-García (2019); this article adds the 2×2×2 factorial analysis, n and means by condition, and an expanded description of the method. **Results:** A one-way ANOVA showed differences among the eight conditions, $F(7, 261) = 2.54$, $p = .015$, $\eta^2 = .06$. The 2×2×2 factorial ANOVA showed a main effect of specificity, F(1, 261) = 9.33, p = .002, partial $\eta^2 = .035$, favoring specific feedback; the main effects of latency, F(1, 261) = 0.97, p = .325, and valence, F(1, 261) = 2.89, p = .090, as well as all interactions (p ≥ .186), were not significant. Mean gain was highest in the specific, appetitive, short-latency condition (4.8 points) and lowest in the generic, appetitive, delayed condition (2.0); the difference was significant in Tukey's test ($p = .041$). **Conclusions:** The data are compatible with specific, timely feedback being associated with larger gains and show no advantage of appetitive stimuli. With one group per condition, differences cannot be attributed to feedback independently of the group.

Keywords: Pedagogical feedback; English as a foreign language; Latency; Quasi-experimental design; Secondary education.

Resumo

Objetivo: Analisar os efeitos da especificidade (específica ou genérica), da latência (curta ou demorada) e da valência (neutra ou apetitiva) do feedback pedagógico sobre o ganho na aprendizagem de adjetivos, comparativos e superlativos em inglês em alunos do oitavo ano. **Métodos:** Estudo quantitativo, quase-experimental, com delineamento fatorial 2×2×2, medição pré-teste e pós-teste e oito grupos intactos de uma instituição pública de San Juan de Pasto (Colômbia), cada um com uma condição. Foram analisados os dados de 269 dos 279 estudantes matriculados. A aprendizagem foi medida com um teste de 25 itens elaborado para o estudo (α = 0,77 no pré-teste e 0,76 no pós-teste). Os dados provêm da tese de doutorado da autora e a comparação global entre condições foi publicada em Torres-García (2019); este artigo incorpora a análise fatorial 2×2×2, o relato de n e médias por condição e uma descrição ampliada do método. **Resultados:** Uma ANOVA de um fator mostrou diferenças entre as oito condições, $F(7, 261) = 2{,}54$, $p = 0{,}015$, $\eta^2 = 0{,}06$. A ANOVA fatorial 2×2×2 mostrou efeito principal da especificidade, F(1, 261) = 9,33, p = 0,002, η^2 parcial = 0,035, a favor do feedback específico; os efeitos principais da latência, F(1, 261) = 0,97, p = 0,325, e da valência, F(1, 261) = 2,89, p = 0,090, assim como todas as interações (p ≥ 0,186), não foram significativos. O ganho médio foi maior na condição específica, apetitiva e curta (4,8 pontos) e menor na genérica, apetitiva e demorada (2,0); a diferença foi significativa

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

https://riied.org

no teste de Tukey (p = 0,041). **Conclusões:** Os dados são compatíveis com a ideia de que o feedback específico e oportuno se associa a maiores ganhos e não mostram vantagem dos estímulos apetitivos. Com um único grupo por condição, as diferenças não podem ser atribuídas ao feedback independentemente do grupo.

Palavras-chave: Feedback pedagógico; Inglês como língua estrangeira; Latência; Delineamento quase-experimental; Ensino Fundamental II.

Introducción

La enseñanza del inglés en educación secundaria constituye un ámbito estratégico para el desarrollo de competencias comunicativas y para la continuidad de las trayectorias académicas de los estudiantes. La formación en una lengua extranjera contribuye a la participación en escenarios culturales, educativos, científicos y laborales mediados por la comunicación global. Sin embargo, el aprendizaje de competencias básicas en inglés no depende únicamente de la exposición a contenidos gramaticales, léxicos o fonológicos. Requiere procesos didácticos que permitan reconocer el desempeño inicial, identificar los errores como oportunidades de aprendizaje y orientar acciones concretas de mejora.

En este contexto, la retroalimentación pedagógica constituye un componente sustantivo de la evaluación formativa. Su función trasciende la comunicación de aciertos o errores, pues aporta información para que el estudiante contraste su desempeño con los criterios esperados, identifique dificultades específicas y modifique progresivamente sus formas de actuación. De acuerdo con Hattie y Timperley (2007), la retroalimentación adquiere mayor valor educativo cuando brinda información sobre la meta de aprendizaje, el nivel actual de desempeño y las acciones necesarias para avanzar hacia el resultado esperado.

La investigación reciente ha confirmado que la evaluación formativa en inglés como lengua extranjera favorece el aprendizaje cuando se integra a la enseñanza cotidiana, se orienta hacia la mejora y permite una participación activa de estudiantes y docentes. La revisión de Zhang et al. (2024), basada en 40 estudios publicados entre 2019 y 2023, identificó que las creencias y percepciones de docentes y estudiantes son decisivas para la aplicación efectiva de la evaluación formativa en contextos K–12 de inglés como lengua extranjera. Asimismo, Rojas Luján (2026) reportó que la retroalimentación formativa en educación secundaria se articula con procesos de corrección, evaluación para el aprendizaje, uso de herramientas digitales, aprendizaje activo y fortalecimiento de la autonomía.

En la enseñanza de lenguas, la retroalimentación correctiva posee una relevancia particular. Esta permite dirigir la atención del estudiante hacia la diferencia entre su producción lingüística y las formas esperadas de la lengua meta. En un estudio con 76 aprendices universitarios de chino como segunda lengua, Zhang et al. (2022) encontraron que las

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

creencias de los estudiantes sobre la retroalimentación correctiva modulan su precisión oral, y sugieren que quién corrige, cuándo corrige y cómo corrige debe ajustarse a necesidades cognitivas y afectivas particulares. Del mismo modo, Ha y Nguyen (2021), con 250 estudiantes y 24 docentes de cuatro colegios públicos de Vietnam, encontraron que unos y otros valoraron la corrección del profesor, aunque los estudiantes también reconocieron la utilidad de la autocorrección y la corrección entre pares cuando reciben orientación para realizarlas.

La efectividad de la retroalimentación no es homogénea. Depende de su contenido, oportunidad, modalidad, foco lingüístico, tarea y recepción por parte del estudiante. En particular, la literatura reciente ha enfatizado que la interacción entre características cognitivas y afectivas del aprendiz puede condicionar el efecto de la retroalimentación correctiva. Liao y Zhang (2022), con 100 estudiantes de secundaria básica en China asignados al azar, mostraron que la memoria de trabajo predijo el efecto de una de las modalidades de retroalimentación correctiva sobre el desarrollo de la segunda lengua.

La latencia, es decir, el intervalo entre la respuesta del estudiante y la información que recibe sobre ella, ha recibido menos atención que el contenido del mensaje en la literatura sobre inglés escolar revisada para este trabajo. La literatura general sobre retroalimentación formativa indica que el efecto de darla de inmediato o con demora depende de la dificultad de la tarea y del nivel del aprendiz (Shute, 2008; Li et al., 2025). En segundas lenguas, la revisión sistemática de Xu y Zeng (2023), con 20 estudios, concluyó que aún no hay una respuesta definitiva sobre el momento óptimo para atender los errores, aunque la retroalimentación inmediata rindió al menos igual que la demorada. Esa falta de consenso justifica probar la latencia de forma experimental en el aula.

La especificidad constituye uno de los atributos centrales de una retroalimentación pedagógica útil, y la propia autora la examinó con estos datos en Torres-García (2019). Un mensaje específico informa qué componente de la respuesta requiere corrección, cuál fue el error, por qué necesita ajustarse y cuál puede ser una vía de mejora. En contraste, una devolución genérica se limita a aprobar o desaprobar el resultado sin precisar el contenido lingüístico o la estrategia pertinente. En estudiantes de secundaria que aprenden inglés como lengua extranjera (53 estudiantes de noveno año en Alemania), Siekmann et al. (2023) encontraron que la percepción de calidad de la retroalimentación y el efecto motivacional atribuido a esta se asociaron con la calidad textual en una segunda medición. El estudio resalta la importancia de proporcionar criterios transparentes, información sobre el desempeño y recomendaciones específicas para mejorar.

La valencia afectiva también debe ser considerada. La retroalimentación puede ser neutra o incorporar estímulos, elogios, recompensas o reforzadores que buscan incrementar la motivación. Sin embargo, el efecto de tales recursos no puede asumirse como uniforme. La evidencia reciente muestra que el compromiso del estudiante con la retroalimentación correctiva comprende dimensiones conductuales, cognitivas y afectivas; por ello, su aprovechamiento

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

depende de cómo el estudiante interpreta y usa la información recibida, no solamente de la intención del docente al emitirla (Liu y Feng, 2023). Esa capacidad de interpretar y usar la información es lo que la literatura denomina alfabetización en retroalimentación (Carless y Winstone, 2023), aunque ese trabajo se refiere a educación superior. En consecuencia, una retroalimentación apreciada como clara, respetuosa y accionable puede favorecer el aprendizaje incluso cuando conserva una valencia afectiva neutra.

Este artículo se basa en la investigación realizada para la tesis doctoral de la autora (Torres-García, 2017). Un primer informe de esos datos, centrado en el contraste global entre condiciones, se publicó en Torres-García (2019), con la misma muestra, el mismo instrumento y la misma tabla de medias. Por tanto, este texto no presenta una recolección nueva. Respecto del artículo de 2019, este texto aporta (a) el análisis factorial 2×2×2, que permite probar efectos principales e interacciones y que no se había realizado; (b) el informe del n y de las medias de preprueba, postprueba y ganancia por condición, junto con los tamaños del efecto; y (c) una descripción ampliada del método y de las limitaciones del diseño. Se declara así para que el lector pueda valorar el aporte y evitar que la misma evidencia se cuente dos veces.

El objetivo de este artículo fue estimar los efectos principales y las interacciones de la especificidad, la latencia y la valencia de la retroalimentación pedagógica sobre la ganancia en el aprendizaje de adjetivos, comparativos y superlativos en inglés en estudiantes de octavo grado, en ocho grupos intactos de una institución pública.

Se plantearon tres preguntas exploratorias. Primera, ¿se asocia la retroalimentación específica con mayores ganancias que la genérica? Segunda, ¿se asocia la retroalimentación de latencia corta con mayores ganancias que la demorada? Tercera, ¿modifica la valencia apetitiva el efecto de las otras dos características? Además, se examinó la interacción entre especificidad y latencia.

Metodología y métodos

Diseño y tipo de estudio

Se realizó un estudio cuantitativo, de alcance explicativo y diseño cuasiexperimental. Se aplicó una estructura factorial 2×2×2 entre grupos, con mediciones preprueba y postprueba. Las variables independientes fueron la especificidad de la retroalimentación (específica o genérica), la latencia (corta o demorada) y la valencia (neutra o apetitiva). La variable dependiente fue la ganancia en el desempeño en adjetivos, comparativos y superlativos en inglés, entre la preprueba y la postprueba.

Cada una de las ocho condiciones se aplicó a un grupo intacto (8-1 a 8-8), de modo que hubo un solo grupo por condición. Por ello el efecto de la condición no puede separarse del efecto del grupo ni del docente que lo atendió (la misma docente, la autora, aplicó la

intervención en los ocho grupos y conocía la condición asignada a cada uno, por lo que sus expectativas no pudieron controlarse). El estudio se clasificó como cuasiexperimental porque los participantes pertenecían a grupos escolares preexistentes. Las ocho condiciones se asignaron al azar a los ocho grupos intactos, una por grupo; los estudiantes, en cambio, no se asignaron aleatoriamente a las condiciones. La «aleatorización» mencionada en Torres-García (2019) se refiere, por tanto, a la asignación de las condiciones a los grupos y no a la de los sujetos.

Contexto y participantes

El estudio se desarrolló en la jornada de la mañana de la Institución Educativa Municipal Ciudad de Pasto, ubicada en San Juan de Pasto, departamento de Nariño, Colombia. Participaron 279 estudiantes de octavo grado, pertenecientes a los grupos 8-1, 8-2, 8-3, 8-4, 8-5, 8-6, 8-7 y 8-8, con edades entre los 12 y los 15 años. En conjunto, el 51 % eran hombres y el 49 % mujeres; por grupo, la proporción de mujeres osciló entre el 35 % (8-8) y el 59 % (8-3). Según la caracterización institucional, alrededor del 70 % de los estudiantes pertenecía a los estratos socioeconómicos 1 y 2, y el 30 % al estrato 3.

Se utilizó un muestreo no probabilístico por conveniencia, determinado por la disponibilidad de ocho grupos intactos. Se consideraron participantes los estudiantes matriculados en octavo grado que asistieron a las sesiones de intervención y contaban con registros utilizables en la preprueba y la postprueba. El análisis inferencial incluyó 269 casos (96.4 %). Los 10 casos restantes quedaron fuera por registros incompletos o por no cumplir los criterios de exclusión definidos de antemano (consentimiento informado firmado por los acudientes, asistencia a todas las sesiones de medición y de intervención, y no pertenecer al programa de inclusión escolar por dificultades severas de aprendizaje). La tesis no desglosa estas pérdidas por grupo, pero los tamaños analizados fueron similares entre condiciones (entre 31 y 36 estudiantes; Tabla 1).

Variables y condiciones de retroalimentación

La estrategia didáctica se organizó en ocho condiciones, derivadas del cruce de tres variables independientes con dos niveles cada una (Tabla 1). Por razones éticas no se incluyó una condición de valencia aversiva.

Tabla 1. Condiciones de retroalimentación pedagógica del diseño factorial

Código	Especificidad	Valencia	Latencia	Grupo	n analizado
EAC	Específica	Apetitiva	Corta	8-1	36
GAC	Genérica	Apetitiva	Corta	8-2	36
EAD	Específica	Apetitiva	Demorada	8-3	31
GAD	Genérica	Apetitiva	Demorada	8-4	36

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

ENC	Específica	Neutra	Corta	8-5	32
GNC	Genérica	Neutra	Corta	8-6	33
END	Específica	Neutra	Demorada	8-7	33
GND	Genérica	Neutra	Demorada	8-8	32

Nota. E = específica; G = genérica; A = apetitiva; N = neutra; C = corta; D = demorada. Fuente: elaboración propia con datos de la investigación.

La retroalimentación específica proporcionó información concreta sobre la respuesta del estudiante, el error identificado o la forma esperada de resolución. La retroalimentación genérica comunicó aprobación o corrección de manera global, sin detallar el componente lingüístico que requería ajuste.

La latencia corta se suministró inmediatamente después de la respuesta del estudiante, y la latencia demorada en un momento posterior (en la corta, el ejercicio se revisaba ítem por ítem, de modo que la información llegaba en un segundo o menos tras cada respuesta; en la demorada, el ejercicio se revisaba completo al terminarlo, sin interrupciones, con un intervalo superior a 10 minutos). La valencia apetitiva se implementó mediante estímulos o reforzadores, y la neutra se limitó a comunicar la información correctiva sin ellos (se entregaba una ficha por cada respuesta correcta, de modo que el número de fichas dependía de los aciertos de cada estudiante; al final de la actividad, la docente contaba las fichas con cada estudiante, las convertía en un puntaje referido a los niveles del Marco Común Europeo de Referencia y le entregaba un diploma con el nivel alcanzado; en la condición neutra no se entregaban fichas y el acierto se reconocía solo con una expresión verbal breve y neutra).

La retroalimentación se aplicó en la clase de inglés de cada grupo, bajo la conducción de la autora, sobre un ejercicio escrito de aplicación de adjetivos, comparativos y superlativos. Al terminarlo, la docente recogía los ejercicios y los redistribuía al azar, de modo que cada estudiante revisaba por escrito el de un compañero según el protocolo de su condición: en la específica, marcaba el error en rojo y escribía al frente la forma correcta; en la genérica, sin marcar la respuesta, encerraba en un círculo el número de la regla transgredida, a partir de las reglas numeradas expuestas en el tablero. Después, los ejercicios se devolvían a sus autores. La intervención comprendió tres sesiones de explicación y ejercitación, precedidas por la preprueba y seguidas por la postprueba, en un total de nueve sesiones de clase distribuidas en tres semanas (15 días hábiles); el número de sesiones diarias y el orden de los contenidos fueron iguales en todos los grupos. La aplicación de cada condición se guio por un protocolo escrito y se documentó con registro fotográfico de las sesiones, de la entrega de fichas y de la codificación de los instrumentos.

Instrumento y evidencia de confiabilidad

El aprendizaje se evaluó con dos versiones de una prueba de 25 ítems de respuesta cerrada, elaborada por la autora para este estudio (no es un instrumento estandarizado), aplicadas como preprueba y postprueba. Los contenidos comprendieron adjetivos calificativos, adjetivos opuestos, género y número del

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

adjetivo, comparativos, superlativos y su aplicación en oraciones y textos breves. El desempeño esperado se vinculó con el establecimiento de comparaciones entre lugares, personas y objetos mediante estructuras y vocabulario básico en inglés. Las respuestas correctas se codificaron con 1 y las incorrectas con 0, por lo que la puntuación varía de 0 a 25. Ambas versiones se construyeron a partir de la misma estructura de prueba (mismos dominios, igual número de ítems y dificultad creciente), con ítems distintos, y se validaron con el mismo procedimiento; la similitud de sus coeficientes de confiabilidad es coherente con esa equivalencia.

Las pruebas se sometieron a una aplicación piloto con 24 estudiantes de grupos distintos de los del estudio principal. La claridad de ítems e instrucciones se valoró con un criterio mínimo de 90 %, y a partir de ella se modificaron las instrucciones y un ítem de la postprueba. Después, cuatro jueces (dos docentes de inglés y dos psicólogos) evaluaron la pertinencia y la claridad de los ajustes.

La consistencia interna, estimada con alfa de Cronbach, fue α = .77 en la preprueba y α = .76 en la postprueba (Tabla 2). Son valores moderados, aceptables para comparar grupos, pero no para decisiones sobre estudiantes individuales. Los coeficientes se estimaron con las respuestas de los estudiantes del estudio principal en cada aplicación.

Tabla 2. Confiabilidad de las pruebas de inglés

Instrumento	Número de ítems	Alfa de Cronbach
Preprueba de inglés	25	.77
Postprueba de inglés	25	.76

Nota. Fuente: elaboración propia con datos de la investigación.

Procedimiento

La investigación se desarrolló en tres fases. En la primera se construyeron, validaron y pilotearon los instrumentos. En la segunda se aplicó la preprueba para identificar el desempeño inicial. En la tercera se implementaron las condiciones de retroalimentación en los ocho grupos y, después, se aplicó la postprueba. La recolección se realizó en el segundo semestre de 2017 (los consentimientos informados de los acudientes se firmaron en octubre de ese año), a lo largo de tres semanas lectivas. Cada estudiante fue identificado con un código que evitó su reconocimiento personal, y los datos se procesaron con IBM SPSS Statistics, versión 20.

Para valorar la equivalencia inicial de los grupos se compararon las puntuaciones de la preprueba entre las ocho condiciones (Tabla 3). Las medias oscilaron entre 9.73 (ENC) y 11.44 (GNC) puntos, con una media global de 10.70. La fuente original no informó una prueba inferencial de esta comparación ni las desviaciones

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

estándar de la preprueba por grupo, por lo que la equivalencia inicial solo puede valorarse descriptivamente; para atenuar esas diferencias, el análisis se centró en la ganancia.

Análisis estadístico

Se calcularon n, media y desviación estándar de la preprueba, la postprueba y la ganancia por condición. La ganancia se definió como la diferencia entre la puntuación de la postprueba y la de la preprueba (diferencia simple). Aunque el plan de análisis de la tesis describía puntuaciones residualizadas, los análisis informados en ella y en este artículo emplean diferencias simples: en cada condición, la diferencia entre las medias de postprueba y preprueba coincide con la ganancia media (Tabla 3).

El primer análisis fue un ANOVA de un factor con las ocho condiciones, con comparaciones múltiples de Tukey; es el análisis global que ya se había publicado, aquí con sus estadísticos completos. Como el diseño es factorial, el análisis principal de este artículo es un ANOVA factorial 2×2×2 entre grupos sobre la ganancia, con efectos principales de especificidad, latencia y valencia y todas sus interacciones (no se empleó ANCOVA, porque requiere los registros individuales; el modelo se estimó con sumas de cuadrados tipo III a partir del n y de las medias de ganancia por condición y de la media cuadrática intragrupo del ANOVA original, procedimiento que reproduce exactamente un modelo factorial completo entre grupos). Se informan *F*, grados de libertad, *p* y η^2 parcial. Se verificaron la normalidad (Shapiro-Wilk por condición) y la homogeneidad de varianzas (Levene); en caso de ANCOVA, también la homogeneidad de las pendientes de regresión. Se fijó $\alpha = .05$. Las comparaciones post hoc se interpretan como exploratorias por su número. La unidad de análisis fue el estudiante; como hubo un solo grupo por condición, no fue posible modelar el efecto del grupo.

Resultados y discusión

El análisis incluyó a 269 de los 279 participantes iniciales (véase el n por condición en la Tabla 1). La Tabla 3 presenta la preprueba, la postprueba y la ganancia por condición. Las medias de preprueba y postprueba proceden de los informes por condición de la tesis, y su diferencia reproduce la ganancia media de cada condición. En la preprueba, las medias por condición se ubicaron entre 9.73 y 11.44 puntos; la condición con mayor ganancia (EAC) partió de 10.89 puntos, cerca de la media global (10.70).

Tabla 3. Preprueba, postprueba y ganancia por condición de retroalimentación

Condición	n	Preprueba *M* (*DE*)	Postprueba *M* (*DE*)	Ganancia *M* (*DE*)
EAC	36	10.89 (–)	15.64 (–)	4.8 (–)
GAC	36	11.15 (–)	13.46 (–)	2.3 (–)

EAD	31	11.07 (–)	14.00 (–)	2.9 (–)
GAD	36	10.69 (–)	12.74 (–)	2.0 (–)
ENC	32	9.73 (–)	14.11 (–)	4.4 (–)
GNC	33	11.44 (–)	14.46 (–)	3.0 (–)
END	33	10.26 (–)	14.50 (–)	4.2 (–)
GND	32	10.31 (–)	13.77 (–)	3.5 (–)

Nota. Puntuaciones sobre 25 ítems. M = media; DE = desviación estándar. Las medias de preprueba y postprueba proceden de los informes por condición de Torres-García (2017), y las de ganancia, del análisis original (con un decimal). –: la fuente no informa la DE por condición; la DE combinada de la ganancia, estimada como la raíz de la media cuadrática del error (Tabla 4), es 3.68. Fuente: elaboración propia con datos de la investigación.

Un ANOVA de un factor mostró diferencias entre las ocho condiciones, $F(7, 261) = 2.54$, $p = .015$, $\eta^2 = .06$ (Tabla 4), un efecto de magnitud pequeña a moderada. En las comparaciones de Tukey, la condición específica, apetitiva y corta superó a la genérica, apetitiva y demorada por 2.71 puntos, $p = .041$, IC 95 % [0.06, 5.36]. Como el límite inferior del intervalo está cerca de cero, la evidencia de esta comparación es débil.

Tabla 4. ANOVA de un factor de la ganancia preprueba-postprueba según condición de retroalimentación

Fuente de variación	Suma de cuadrados	GI	Media cuadrática	*F*	*p*
Entre grupos	240.52	7	34.36	2.54	.015
Dentro de grupos	3529.43	261	13.52		
Total	3769.95	268			

Nota. $\eta^2 = .064$. Fuente: elaboración propia con datos de la investigación.

El ANOVA factorial 2×2×2 (sumas de cuadrados tipo III; Tabla 5) mostró un efecto principal significativo de la especificidad, $F(1, 261) = 9.33$, $p = .002$, η^2 parcial = .035: la retroalimentación específica produjo una ganancia media mayor que la genérica (4.08 frente a 2.70 puntos, medias no ponderadas). El efecto principal de la latencia no fue significativo, $F(1, 261) = 0.97$, $p = .325$, η^2 parcial = .004, ni el de la valencia, $F(1, 261) = 2.89$, $p = .090$, η^2 parcial = .011. Ninguna interacción alcanzó significación: especificidad × latencia, $F(1, 261) = 1.39$, $p = .239$, η^2 parcial = .005; especificidad × valencia, $F(1, 261) = 0.44$, $p = .509$, η^2 parcial = .002; latencia × valencia, $F(1, 261) = 1.76$, $p = .186$, η^2 parcial = .007; y especificidad × latencia × valencia, $F(1, 261) = 0.30$, $p = .586$, η^2 parcial = .001. Por tanto, no se sostiene que la especificidad y la latencia interactuaran. Los supuestos

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

se cumplieron: la prueba de Shapiro-Wilk no fue significativa en ninguna condición (p entre .156 y .713) y la prueba de Levene indicó homogeneidad de varianzas, F(7, 261) = 1.22, p = .291.

Tabla 5. ANOVA factorial 2×2×2 de la ganancia según especificidad, latencia y valencia

Fuente de variación	Suma de cuadrados	gl	Media cuadrática	*F*	*p*	η^2 parcial
Especificidad (E)	126.15	1	126.15	9.33	.002	.035
Latencia (L)	13.17	1	13.17	0.97	.325	.004
Valencia (V)	39.05	1	39.05	2.89	.090	.011
E × L	18.85	1	18.85	1.39	.239	.005
E × V	5.92	1	5.92	0.44	.509	.002
L × V	23.81	1	23.81	1.76	.186	.007
E × L × V	4.02	1	4.02	0.30	.586	.001
Error	3529.43	261	13.52			

Nota. Sumas de cuadrados tipo III, calculadas a partir del n y de las medias de ganancia por condición y de la media cuadrática intragrupo del ANOVA de un factor (Tabla 4). Fuente: elaboración propia con datos de la investigación.

Descriptivamente, la ganancia media fue mayor con retroalimentación específica y corta (4.6 puntos) que con específica y demorada (3.6), mientras que la genérica obtuvo 2.7 puntos con ambas latencias. Por valencia, la media de la retroalimentación neutra (3.8) fue superior a la de la apetitiva (3.0) (Tabla 6). Estas diferencias se interpretan solo en la medida en que las respalde el modelo factorial.

Tabla 6. Ganancias medias por condición y promedios por componente de la retroalimentación

Valencia	Específica demorada	Específica corta	Genérica demorada	Genérica corta	Promedio por valencia
Apetitiva	2.9	4.8	2.0	2.3	3.0
Neutra	4.2	4.4	3.5	3.0	3.8
Promedio por especificidad y latencia	3.6	4.6	2.7	2.7	

Nota. Los promedios son medias no ponderadas de las medias de celda. Se verificó que corresponden a este cálculo: con ponderación por n, el promedio de la genérica corta sería 2.6 y no 2.7. Fuente: elaboración propia con datos de la investigación.

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

Las diferencias entre condiciones apuntan en la dirección esperada para la especificidad y la latencia, pero son de magnitud pequeña a moderada y se apoyan en un contraste puntual frágil (específica, apetitiva y corta frente a genérica, apetitiva y demorada). Según el modelo factorial, solo la especificidad tuvo un efecto principal significativo, de magnitud pequeña (η^2 parcial = .035); ni la latencia ni la interacción entre especificidad y latencia alcanzaron significación. Aunque, descriptivamente, la ventaja de la retroalimentación específica fue mayor con entrega inmediata (4.6 frente a 3.6 puntos) y la genérica rindió lo mismo, 2.7 puntos, con latencia corta y demorada, la falta de significación de la interacción impide interpretar este patrón como un efecto combinado.

Este patrón es compatible con lo que encontraron Xu y Zeng (2023), cuya revisión no halló ventaja de la demora, y con la idea de que el efecto del momento depende de la tarea (Shute, 2008). Que el mensaje específico rinda más cuando llega de inmediato admite una explicación plausible, la de asociar la información con la respuesta emitida y poder corregirla en el mismo ciclo, pero el estudio no registró cómo atendieron los estudiantes los mensajes; es una interpretación y no un resultado. En cuanto a la especificidad, el hallazgo coincide con lo observado por Siekmann et al. (2023) en estudiantes alemanes de noveno año, donde la calidad percibida de la retroalimentación se asoció con la calidad de los textos, aunque en un diseño y con una variable de resultado distintos.

La valencia apetitiva no mostró ventaja, y la media de la retroalimentación neutra fue descriptivamente mayor. En Torres-García (2019) se propuso que las fichas pudieron no tener propiedades apetitivas funcionales para los estudiantes. La explicación es posible, pero este diseño no la probó: no se midió qué consecuencias eran apetitivas para cada estudiante. Por eso la ausencia de ventaja no permite concluir que los estímulos carezcan de efecto en general, sino que estos estímulos, entregados de esta manera, no mostraron ventaja en esta muestra. Los trabajos sobre compromiso con la retroalimentación (Liu y Feng, 2023) y sobre las creencias de los estudiantes (Zhang et al., 2022) sugieren que el valor de un estímulo depende de cómo lo interpreta quien lo recibe.

Para la práctica, los datos sugieren que, cuando el tiempo es limitado, conviene priorizar mensajes breves que indiquen qué parte de la respuesta debe ajustarse (vocabulario, uso del adjetivo, forma comparativa o superlativa, concordancia) y darlos durante la actividad o inmediatamente después. Es una orientación provisional, porque procede de tres sesiones sobre un solo contenido y no debe tomarse como regla. Ha y Nguyen (2021) añaden que los estudiantes valoran también la autocorrección y la corrección entre pares cuando reciben entrenamiento, por lo que combinar modalidades es una vía por explorar.

Las limitaciones son varias. Primera, hubo un solo grupo por condición, de modo que condición, grupo y docente están confundidos, y las diferencias no pueden atribuirse a la retroalimentación con independencia del grupo. Segunda, no hubo asignación aleatoria de estudiantes (las condiciones se asignaron al azar a grupos intactos, no a estudiantes) y la equivalencia inicial debe mostrarse con los datos de la preprueba. Tercera, la intervención fue breve (tres sesiones de

¿Cómo citar el artículo?
Torres-García, Z. J. (2026). Efecto de la retroalimentación pedagógica en el aprendizaje de gramática inglesa en octavo grado. *RIIED, número* 10, 1-15.

intervención, dentro de nueve sesiones de clase en tres semanas), con un solo contenido, en una sola institución y sin medición de seguimiento. Cuarta, el instrumento fue elaborado para el estudio y su consistencia interna es moderada. Quinta, se perdieron 10 casos (3.6 %). Sexta, las comparaciones post hoc son numerosas y la que sostiene la lectura tiene un intervalo de confianza que casi toca el cero. Séptima, los datos se recogieron en 2017 y se reanalizan ahora, por lo que su vigencia debe valorarse con esa fecha.

Investigaciones futuras podrían replicar el diseño con varios grupos por condición asignados al azar, ampliar los contenidos evaluados, incluir seguimiento y comparar modalidades oral, escrita y mediada por tecnologías digitales.

Conclusiones

Entre las ocho condiciones se observaron diferencias en la ganancia de aprendizaje de adjetivos, comparativos y superlativos ($F(7, 261) = 2.54$, $p = .015$, $\eta^2 = .06$). El modelo factorial indicó que solo la especificidad tuvo un efecto principal significativo ($F(1, 261) = 9.33$, $p = .002$, η^2 parcial $= .035$): la retroalimentación específica se asoció con mayores ganancias que la genérica (4.1 frente a 2.7 puntos), mientras que la latencia, la valencia y las interacciones no fueron significativas.

La retroalimentación específica y de latencia corta mostró la mayor ganancia media entre los cuatro pares de especificidad y latencia (4.6 puntos), mientras que la genérica obtuvo 2.7 con ambas latencias. Esto es compatible con una ventaja de la información específica y oportuna, aunque la latencia no tuvo un efecto principal significativo y la prueba de interacción no respaldó un efecto combinado.

La valencia apetitiva no mostró ventaja: la media de la retroalimentación neutra (3.8) superó, en términos descriptivos, a la de la apetitiva (3.0).

Dadas las limitaciones del diseño (un grupo por condición, intervención breve, un solo contenido y una institución), los resultados orientan, pero no establecen una relación causal ni se extienden a otras habilidades. Se recomienda replicar el estudio con varios grupos por condición asignados al azar, otras habilidades lingüísticas y mediciones de seguimiento.

Referencias bibliográficas

Carless, D., & Winstone, N. E. (2023). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education, 28*(1), 150–163. https://doi.org/10.1080/13562517.2020.1782372

Ha, X. V., & Nguyen, L. T. (2021). Targets and sources of oral corrective feedback in English as a foreign language classrooms: Are students' and

teachers' beliefs aligned? *Frontiers in Psychology, 12*, Article 697160. https://doi.org/10.3389/fpsyg.2021.697160

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Li, S., Ou, L., & Lee, I. (2025). The timing of corrective feedback in second language learning. *Language Teaching, 58*(1), 1–17. https://doi.org/10.1017/s0261444824000478

Liao, Y., & Zhang, W. (2022). Corrective feedback, individual differences in working memory, and L2 development. *Frontiers in Psychology, 13*, Article 811748. https://doi.org/10.3389/fpsyg.2022.811748

Liu, H., & Feng, M. (2023). The role of learner engagement with corrective feedback in EFL/ESL classrooms. *Frontiers in Psychology, 14*, Article 1118467. https://doi.org/10.3389/fpsyg.2023.1118467

Rojas Luján, N. J. (2026). Retroalimentación formativa en la enseñanza del inglés en educación secundaria: Una revisión sistemática (2021–2025). *Revista InveCom, 6*(3), e603043. https://doi.org/10.5281/zenodo.17612944

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Siekmann, L., Parr, J., Van Ophuysen, S., & Busse, V. (2023). Text quality and changing perceptions of teacher feedback and affective-motivational variables: A study with secondary EFL students. *Frontiers in Education, 8*, Article 1171914. https://doi.org/10.3389/feduc.2023.1171914

Torres-García, Z. J. (2017). *Efectos de variaciones de la retroalimentación pedagógica en el aprendizaje de competencias básicas de inglés* [Tesis doctoral, Universidad de Baja California, Tepic; México].

Torres-García, Z. J. (2019). Efectos de variaciones en la retroalimentación pedagógica en el aprendizaje de inglés. *Cuadernos de Lingüística Hispánica*, (34), 215–232. https://doi.org/10.19053/0121053X.n34.2019.9722

Xu, M., & Zeng, S. (2023). Optimal timing of treatment for errors in second language learning – A systematic review of corrective feedback timing. *Frontiers in Psychology, 14*, Article 1026174. https://doi.org/10.3389/fpsyg.2023.1026174

Zhang, H., Ge, S., & Mohd Saad, M. R. (2024). Formative assessment in K-12 English as a foreign language education: A systematic review. *Heliyon, 10*, Article e31367. https://doi.org/10.1016/j.heliyon.2024.e31367

Zhang, J., Cao, X., & Zheng, N. (2022). How learners' corrective feedback beliefs modulate their oral accuracy: A comparative study on high- and low-accuracy learners of Chinese as a second language. *Frontiers in Psychology, 13*,

Article 869468.
https://doi.org/10.3389/fpsyg.2022.869468

Contribución autoral:

Zoila Josefina Torres García: Conceptualización, Metodología, Investigación, Análisis formal, Redacción, revisión y edición.

Publicación previa: Los datos de este artículo proceden de la tesis doctoral de la autora (Torres-García, 2017) y su contraste global entre condiciones se publicó en Torres-García (2019).

Conflicto de intereses

La autora declara no tener conflictos de interés.