



**ANÁLISIS DE ERRORES MORFOSINTÁCTICOS EN EL PARADIGMA
NOMINAL EN NIVELES A2, B1 Y B2 EN UN CORPUS DE APRENDIENTES
INDIOS**

Presentado por

Amrit Goyal

Directora:

Andrea Marcela Baquero Caldas

Mgtr. en Enseñanza de Español como Lengua Extranjera y Segunda Lengua

**MAESTRÍA EN ENSEÑANZA DE ESPAÑOL COMO LENGUA EXTRANJERA Y
SEGUNDA LENGUA**

INSTITUTO CARO Y CUERVO

BOGOTÁ, 2024

Agradecimientos

A Rashmi, Vinod y Pushkar (mi familia) por intentar de entender mi pasión y apoyarme,

A todos mis profesores (mi familia colombiana) que me amaron tanto,

A mis parceros por animarme durante momentos difíciles,

A Claudia por guiarme de forma divertida,

A Andrea por ser y estar,

Y a mí por intentar.

AUTORIZACIÓN DEL AUTOR PARA CONSULTA Y PUBLICACIÓN ELECTRÓNICA DEL TRABAJO DE GRADO

BIBLIOTECA JOSÉ MANUEL RIVAS SACCONI

INFORMACION DEL TRABAJO DE GRADO

1. *Trabajo de grado requisito para optar al título de: Maestría en enseñanza de español como lengua extranjera y segunda lengua.*

2. **Título del trabajo de grado:**

Análisis de errores morfosintácticos en el paradigma nominal en niveles A2, B1 y B2 en un corpus de aprendientes indios

3. **Autoriza la consulta y publicación electrónica del trabajo de grado:**

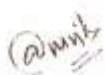
Sí autorizo ☒ , No autorizo ☐ , a la biblioteca José Manuel Rivas Sacconi del Instituto Caro y Cuervo para que con fines académicos:

- Ponga el contenido de este trabajo a disposición de los usuarios en la biblioteca digital Palabra, así como en redes de información del país y del exterior, con las cuales tenga convenio la Facultad Seminario Andrés Bello y el Instituto Caro y Cuervo.
- Permita la consulta a los usuarios interesados en el contenido de este trabajo, para usos de finalidad académica, ya sea formato impreso, CD-ROM o digital desde Internet.
- Socialice la producción intelectual de los egresados de las Maestrías del Instituto Caro y Cuervo con la comunidad académica en general.
- Todos los usos, que tengan finalidad académica; de manera especial la divulgación a través de redes de información académica.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, **“Los derechos morales sobre el trabajo son propiedad de los autores”**, los cuales son irrenunciables, imprescriptibles, inembargables e inalienables. Atendiendo lo anterior, siempre que se consulte la obra, mediante cita bibliográfica se debe dar crédito al trabajo y a su autor.

4. **Identificación del autor**

Firma:



Nombre completo: Amrit Goyal

Documento de identidad: Pasaporte: T5228852

DESCRIPCIÓN TRABAJO DE GRADO

AUTOR

Apellidos	Nombres
Goyal	Amrit

DIRECTOR (ES)

Apellidos	Nombres
Baquero Caldas	Andrea Marcela

TRABAJO PARA OPTAR POR EL TÍTULO DE: Maestría en enseñanza de español como lengua extranjera

TÍTULO DEL TRABAJO DE GRADO:

Análisis de errores morfosintácticos en el paradigma nominal en niveles A2, B1 y B2 en un corpus de aprendientes indios

NOMBRE DEL PROGRAMA ACADÉMICO:

Maestría en enseñanza de español como lengua extranjera (2022-2024)

CIUDAD: Bogotá AÑO DE PRESENTACIÓN DEL TRABAJO: 2024

NÚMERO DE PÁGINAS: 142

TIPO DE ILUSTRACIONES: Ilustraciones ____ Mapas ____ Retratos ____ Tablas, gráficos y diagramas _83_ Planos ____ Láminas ____ Fotografías ____

MATERIAL ANEXO (Vídeo, audio, multimedia):

Duración del audiovisual: _____ Minutos.

Otro. ¿Cuál? _____

Sistema: Americano NTSC _____ Europeo PAL _____ SECAM _____

PREMIO O DISTINCIÓN (En caso de ser Laureadas o tener una mención especial):

DESCRIPTORES O PALABRAS CLAVES: Son los términos que definen los temas que identifican el contenido. *(En caso de duda para designar estos descriptores, se recomienda consultar a la dirección de biblioteca en el correo electrónico biblioteca@caroycuervo.gov.co):*

ESPAÑOL

Corpus de aprendientes indios, Errores morfosintácticos, Español como lengua extranjera, Necesidades de aprendientes indios, Lengua materna hindi.

INGLÉS

Corpus of Indian learners, Morphosyntactic errors, Spanish as a foreign language, Needs of Indian learners, Hindi as the mother tongue.

RESUMEN DEL CONTENIDO

Esta investigación tenía el objetivo de analizar los errores morfosintácticos en el paradigma nominal de un corpus de aprendientes de español hindi hablantes de niveles A2, B1 y B2. En total se eligieron 59 muestras escritas de 23 participantes con lengua materna hindi, de un corpus de 40 estudiantes con diferentes lenguas maternas de India. El corpus tenía textos descriptivos, narrativos y argumentativos e hizo parte del Corpus de aprendientes de español como lenguas extranjera y segunda lengua (CAELE/2) del Instituto Caro y Cuervo. La investigación se basó en una metodología mixta y se fundamentó en el análisis contrastivo, análisis de errores y la metodología de la lingüística de corpus. Se adaptó una taxonomía desde el criterio lingüístico y etiológico y, se hizo y se actualizó el análisis contrastivo entre hindi y español en el paradigma nominal. Los resultados del análisis revelaron que la interlengua de los estudiantes se permeaba de la interferencia de hindi en todos los niveles. En la misma línea, el hallazgo principal fue la fosilización de la omisión de artículo y la confusión en el concepto de la concordancia en nivel B2. Al final, se hicieron recomendaciones con base en los resultados revelados en este estudio.

Abstract

This research had the objective of analyzing morphosyntactic errors in the nominal paradigm in a corpus of Hindi-speaking, Spanish language learners in levels A2, B1 and B2. In total 59 written samples of 23 participants with mother tongue Hindi were chosen from a corpus of 40 students with different mother languages from India. The corpus had descriptive, narrative and argumentative texts and was part of the Corpus of learners of Spanish as a foreign and second language (CAELE/2) of the Institute Caro y Cuervo. The research was based on a mixed methodology and was theoretically founded on contrastive analysis, error analysis and the methodology of corpus linguistics. A taxonomy was adapted from the linguistic and etiological criteria in the nominal paradigm. The contrastive analysis between Hindi and Spanish in the nominal system was brought about and updated. The results of the analysis revealed that the students' interlanguage was permeated with Hindi interference at all levels. In addition, the main finding was the fossilization of article omission and confusion in the concept of grammatical concordance in level B2. In the end, some recommendations were made, based on the results revealed in this study.

Tabla de contenido

Introducción.....	11
1. Planteamiento del problema de investigación.....	13
2. Objetivos	15
2.1. Objetivo general	15
2.2 Objetivos específicos	15
3. Estado del arte	16
3.1 Análisis contrastivo en lenguas geográfica y lingüísticamente cercanas al español	16
3.2 Análisis contrastivo en lenguas geográfica y lingüísticamente lejanas al español.....	20
3.2.1 Análisis de errores en India	24
4. Justificación	27
5. Marco teórico.....	29
5.1.1 El contraste entre el sistema morfosintáctico de español e hindi con el paradigma nominal.....	32
5.1.1.1 Flexión	33
5.1.1.2 Sustantivos	34
5.1.1.3 Adjetivos.....	39
5.1.1.4 Artículos	42
5.1.1.5 Determinantes.....	48
5.1.1.6 Orden sintáctico.....	63
5.2.1 El concepto del error.....	65
5.2.1.1 El error considerado como síntoma de fracaso y síntoma del progreso.....	66
5.3 La interlengua.....	68
5.4 Las taxonomías para el análisis de los errores.....	70
5.5 La lingüística de corpus	77
6. Marco metodológico	79
6.1 Diseño, alcance y paradigma	79
6.2 Muestra y población.....	81
6.3 Instrumentos.....	82
6.3.1 Cuestionario (Recopilación de datos en la población india).....	82
6.3.2 Taxonomía para la clasificación de errores	85
6.4 Herramienta para el análisis de datos (UAM corpus tool).....	87
6.5 Fases de la investigación	88
6.5.1 Recopilación de corpus	88
6.5.2 Sistematización de datos	89

6.5.3 Validación del instrumento.....	90
6.5.4 Análisis de metadatos.....	91
6.5.5 Criterio para el análisis de datos.....	91
6.6 Consideraciones éticas	92
7. Resultados	93
7.1 Nivel A2.....	95
7.1.1Criterio lingüístico.....	95
7.1.2 Criterio etiológico	101
7.2 Nivel B1	105
7.2.1 Criterio lingüístico.....	105
7.2.2 Criterio etiológico	111
7.3 Nivel B2	114
7.3.1 Criterio lingüístico.....	114
7.3.2 Criterio etiológico	122
8. Discusión	130
9. Conclusiones	136
10.Referencias.....	141
11. Anexos	146

Lista de tablas

Tabla 1: La flexión de género y número en adjetivos en español	33
Tabla 2: El sistema de números en los sustantivos	34
Tabla 3: Declinaciones de ladka (muchacho).....	36
Tabla 4: Declinaciones de beti (hija).....	36
Tabla 5: Declinaciones de geet (canción).....	37
Tabla 6: Declinaciones de baat (palabra)	37
Tabla 7: Declinaciones de aadmi (hombre).....	38
Tabla 8: Declinación de maataa (madre).....	38
Tabla 9: Flexión de género y número en el adjetivo ‘blanco’ y ‘bueno’	40
Tabla 10: Declinación de lamba (alto, largo): cuya forma termina en -a.....	41
Tabla 11: Declinación de mushkil (difícil): cuya forma no termina en -a	41
Tabla 12: Artículos definidos en español.....	42
Tabla 13: Artículos indefinidos en español	43
Tabla 14: Artículos en hindi y español.....	48
Tabla 15: Los demostrativos en hindi y español	49
Tabla 16: Los posesivos antepuestos en español.....	51
Tabla 17: Los posesivos pospuestos en español.....	51
Tabla 18: Comparación de los indefinidos en hindi y español.....	54
Tabla 19: Números cardinales en español e hindi	56

Tabla 20: Agregativos en hindi	58
Tabla 21: Números ordinales en hindi y español	59
Tabla 22: Declinaciones de número ordinal dusra (segundo/a)	61
Tabla 23: Relativos determinantes en español	61
Tabla 24: Declinaciones de relativo jitna en hindi	62
Tabla 25: Declinaciones de relativo jiska en hindi.....	62
Tabla 26: Determinantes relativos en hindi y español	62
Tabla 27: Ferreira & Elejalde (2020).....	71
Tabla 28: Tipología de errores propuesto por Santos Gargallo (1993).....	72
Tabla 29: Tipología de errores de Alexopoulou (2005).....	72
Tabla 30: Tipología de errores propuesta por Duraó (2005).....	73
Tabla 31: Taxonomía para la clasificación de errores proyecto Fondecyt n.º 1110812 y n.º 1180974	74
Tabla 32: Propuesta de taxonomía con criterio etiológico (Ferreira y Elejalde, 2020).....	75
Tabla 33: Errores de transferencia de estructuras gramaticales	76
Tabla 34: Preguntas para recoger las muestras escritas	84
Tabla 35: Taxonomía general para el primer etiquetado de errores.....	85
Tabla 36: Taxonomía adaptada de Ferreira y Elejalde (2020)	86
Tabla 37: Codificación para el perfil del estudiante.....	89
Tabla 38: Codificación de la muestra escrita	89
Tabla 39: Tipos de archivos en el corpus CAELE/2	89
Tabla 40: Análisis preliminar de errores de nivel A2 con 19 muestras.....	93
Tabla 41: Análisis preliminar de errores de nivel B1 con 14 muestras.....	93
Tabla 42: Análisis preliminar de errores de nivel B2 con 27 muestras.....	94
Tabla 43: Muestras de los errores en la categoría gramatical en nivel A2.....	96
Tabla 44: Muestras de los errores en los artículos en nivel A2.....	98
Tabla 45: Muestras de los errores en el aspecto sintáctico en nivel A2	99
Tabla 46: Muestras de los errores de la morfología en nivel A2.....	100
Tabla 47: Muestras de los errores desde el criterio etiológico (interlingüístico) en nivel A2.....	102
Tabla 48: Muestras de cambio de género de español a hindi	103
Tabla 49: Muestras de los errores desde el criterio etiológico (interlingüístico) en nivel A2.....	104
Tabla 50: Muestras de los errores menos frecuentes en nivel B1	107
Tabla 51: Muestras de los errores en los artículos	108
Tabla 52: Muestras de los errores en el aspecto sintáctico de nivel B1	110
Tabla 53: Muestras de los errores de la morfología en nivel B1.....	111
Tabla 54: Muestras de los errores de la interferencia de hindi en nivel B1	112
Tabla 55: Muestras de los errores de la interferencia de español en nivel B1	113
Tabla 56: Muestras de los errores en la categoría gramatical en nivel B2	115
Tabla 57: Muestras de los errores en los artículos en nivel B2	117
Tabla 58: Muestras de los errores desde el aspecto sintáctico en nivel B2	119
Tabla 59: Muestras de los errores de la morfología en nivel B2.....	120
Tabla 60: Muestras de los errores desde la transferencia de hindi en nivel B2.....	122
Tabla 61: Muestras de los errores intralingüísticos en nivel B2	124

Lista de figuras

Figura 1: Representación gráfica de la relación entre hindi y español. Adaptado del libro: Gramática de hindi y urdu para castellanohablantes (Díaz, 2023).....	32
Figura 2: Contraste del género del sustantivo libro en español e hindi, ilustración propia.....	39
Figura 3: Los errores como parte de la Interlengua. Tomado de León (2019).....	69
Figura 4: Esquema del diseño explicativo secuencial adaptado de (Hernández y Mendoza, 2014)	80
Figura 5: Errores lingüísticos en nivel A2.....	95
Figura 6: Errores en las categorías gramaticales de nivel A2	96
Figura 7: Errores de los artículos en nivel A2.....	98
Figura 8: Errores en el aspecto sintáctico de nivel A2	99
Figura 9: Errores desde el criterio etiológico en nivel A2	101
Figura 10: Errores lingüísticos en nivel B1.....	106
Figura 11: Errores en categoría gramatical de nivel B2.....	107
Figura 12: Errores de los artículos en nivel B1	108
Figura 13: Errores sintácticos en nivel B1	110
Figura 14: Errores desde el criterio etiológico en nivel B1	111
Figura 15: Errores lingüísticos en nivel B2.....	115
Figura 16: Errores en la categoría gramatical en nivel B2	115
Figura 17: Errores en los artículos definidos en nivel B2	117
Figura 18: Errores desde el aspecto sintáctico en nivel B2	119
Figura 19: Errores desde el criterio etiológico en nivel B2.....	122
Figura 20: Transición de los errores lingüísticos de nivel A2 a B2	126
Figura 21: Transición de los errores en la categoría gramatical desde nivel A2 a nivel B2	127
Figura 22: Transición de los errores interlingüísticos e intralingüísticos de nivel A2 a B2.....	129

Introducción

La India es un país plurilingüístico. Según el censo de la India, 2011, hay alrededor de 780 lenguas con 22 lenguas oficiales y otras lenguas no oficiales, dentro de las cuales también hay algunas lenguas extranjeras. Dado que el país fue una colonia de muchos países, tiene influencias lingüísticas y culturales de diversos idiomas con raíces lingüísticas diferentes. Uno de estos idiomas fue el inglés, porque la Gran Bretaña colonizó a India y esto implicó la introducción del inglés en el sistema educativo y político a nivel nacional. Por eso, actualmente el país sigue un sistema educativo bilingüe donde el inglés es el segundo idioma enseñado desde la primaria en muchos colegios nacionales e internacionales. También, se enseña un tercer y cuarto idioma como el sánscrito, pero a pesar de ofrecer el aprendizaje de este idioma en la actualidad, por la globalización y la internacionalización del país, las lenguas que gozan de más prestigio son las lenguas extranjeras como el francés y el alemán.

La enseñanza del *español como lengua extranjera* (ELE) llegó en los años cincuenta cuando “en 1956, el gobierno indio becó a dos estudiantes que fueron enviados a España en un programa destinado a mandar alumnos indios a países donde se hablan las lenguas más importantes del planeta” (Pujol Riembau, 2012). Luego, en 1968 el gobierno indio lanzó la política nacional de educación (National Policy on Education) que dio énfasis a la enseñanza de los idiomas extranjeros. Así llegaron varios institutos de lenguas extranjeras (LE) como el *Goethe Institut* de Alemania para enseñar alemán y la *Alliance Francaise* para enseñar francés. Bharti (2019) indica cómo los años setenta fueron los más significativos para fomentar la enseñanza de LE, también porque se fundaron las universidades: *Jawaharlal Nehru University* (JNU) en Nueva Delhi y el *Central Institute of English and Foreign Languages* (CIEFL) en Hyderabad. En la actualidad hay más de veintiocho (28) universidades públicas y privadas, alrededor de sesenta y cuatro (64)

institutos y otros colegios privados que enseñan español en todo el país. Sin embargo, todavía hay problemas con la enseñanza de ELE no regulada, debido a la escasez de docentes calificados, material didáctico y docentes indios con formación (Bharti, 2019).

En este orden de ideas, para guiar al lector y desarrollar el tema de esta investigación, este trabajo inicia su recorrido con una búsqueda de las tendencias investigativas en el mundo, después entra en una indagación más precisa en Asia y en la India. Luego, se exponen los objetivos generales y específicos del proyecto. Asimismo, la justificación se da desde la relevancia social, la implicación práctica y la utilidad metodológica para explicar el porqué de este proyecto de investigación.

En el marco teórico se expondrán las teorías del análisis contrastivo, análisis de errores y la interlengua de los aprendientes. Bajo el análisis contrastivo se indagará sobre el contraste morfosintáctico en el sistema nominal de hindi y español. Posteriormente, se presentarán varias taxonomías relevantes para este proyecto y la adaptación de ellas a partir de los hallazgos preliminares en una exploración inicial del corpus. Luego, se hará el análisis al nivel morfosintáctico de la lengua a partir de la taxonomía adaptada y el paradigma morfosintáctico.

En cuanto al marco metodológico, se describirá el instrumento para la realización de la recopilación del corpus de muestras escritas para aprendientes de español en India. Después, se definirá el segundo instrumento, la taxonomía adaptada para este proyecto con los criterios de selección de las categorías de análisis. Se describirá el *software UAM corpus tool* usado para la categorización de los errores y la sistematización de los datos (muestras escritas).

A continuación, en los resultados se describirán las frecuencias de los errores según la taxonomía adaptada. Se mencionarán las categorías como la omisión de artículo y la frecuencia de los errores más comunes que enfrentan los estudiantes desde la teoría de

análisis contrastivo y análisis de errores. En último lugar, se elaborarán algunas conclusiones con recomendaciones para los docentes, como tener conocimiento lingüístico básico de hindi y las necesidades de los estudiantes hindi hablantes para didactizar las clases, entre otras. Asimismo, se mencionarán las limitaciones y las posibilidades de futuros proyectos que se pueden desprender de esta investigación.

1. Planteamiento del problema de investigación

En primer lugar, el sistema de educación en la India todavía está basado en las necesidades de los aprendientes no originarios indios. Esto se puede ejemplificar porque en el currículo de la mayoría de las escuelas, universidades e institutos se utilizan manuales y material didáctico hecho para europeos o estadounidenses. En segundo lugar, los docentes originarios de la India no tienen formación suficiente en ELE por la falta de universidades que ofrezcan esta capacitación (Bharti, 2019). En tercer lugar, los docentes extranjeros no saben la lengua materna de los indios y, asimismo, no logran aprovechar la riqueza de los conocimientos lingüísticos ya desarrollados por ellos (Bharti, 2019). Estos vacíos en la formación de los profesores y el poco interés en las investigaciones en ELE en la India nos dirigen a un problema común que es la falta de análisis de las necesidades de los estudiantes indios y de su interlengua.

Ahora bien, durante el aprendizaje de un idioma, el estudiante crea diferentes esquemas correspondientes al sistema lingüístico de una lengua, los comprueba y los actualiza dependiendo del rendimiento de la hipótesis en varios contextos (Ferreiro, 2002). Este sistema lingüístico del estudiante de una segunda lengua o lengua extranjera, en cada uno de los estadios sucesivos de aprendizaje por los que pasa, se llama *interlengua* (Selinker, 1969,1972) y la base de la actualización de esta interlengua son los errores. Estos errores se pueden ver como una desviación negativa desde la lengua meta o una oportunidad para

entender mejor las necesidades de los aprendientes. El error bajo esta perspectiva de mejora es una oportunidad de enseñanza y aprendizaje tanto para los docentes como para los aprendientes. De esta forma, el análisis de errores ayuda a los docentes investigadores a anticipar los errores y didactizarlos según las necesidades que surjan de este análisis. Por consiguiente, se puede evitar la fosilización de los errores y actualizar la interlengua en cada nivel de la lengua.

Azevedo (1980) deja constancia de que los estudios de interlengua (IL) ayudan a entender cómo los estudiantes interpretan e internalizan las reglas de la lengua meta. Así, el análisis de errores puede ser una buena herramienta para mejorar la comprensión de la interlengua y trabajar el error desde una perspectiva optimista. No obstante, durante la búsqueda de los antecedentes, se encontró que hay un vacío en el análisis de errores para los estudiantes indios. Tampoco hay un corpus de español designado para ellos a pesar de ser una población grande.

Al analizar esta historia de ELE en la India y la enseñanza de ELE no regulada, en este trabajo, se identificó baja producción investigativa sobre las necesidades de los estudiantes indios y poco material didáctico/adaptado para esta población. Con el propósito de contribuir a llenar este vacío, esta investigación busca analizar los errores morfosintácticos en la población de hindi hablantes para luego entender sus necesidades de aprendizaje. Para ello, se hizo una recopilación de las muestras escritas de los estudiantes indios de diferentes lenguas maternas con diferentes niveles de español. Posteriormente, se delimitó el análisis de errores a niveles de lengua con máximas muestras representativas y se analizaron a nivel morfosintáctico en el paradigma nominal con estudiantes de lengua materna hindi. Este análisis se podría explotar para la creación de material didáctico para los estudiantes con lengua materna hindi y también para estudiar la interlengua de los estudiantes indios en posibles futuras investigaciones.

Según el planteamiento del problema, se propone la siguiente pregunta de investigación: ¿Cuáles son los errores morfosintácticos en el paradigma nominal en las producciones escritas de aprendientes de español de nivel A2, B1 y B2, con lengua materna hindi?

2. Objetivos

2.1. Objetivo general

- Analizar los errores morfosintácticos en el paradigma nominal de estudiantes de español hindi hablantes en niveles A2, B1 y B2, en muestras escritas del corpus CAELE/2, desde la perspectiva de análisis contrastivo (AC) y análisis de errores (AE).

2.2 Objetivos específicos

- Describir las diferencias y similitudes en el paradigma nominal de hindi y español.
- Adaptar una taxonomía para etiquetar y clasificar los errores morfosintácticos desde el criterio lingüístico y etiológico.
- Identificar la frecuencia de los errores morfosintácticos en el nivel A2, B1 y B2 en el paradigma nominal en las muestras escritas de aprendientes indios que hablan hindi como lengua materna, disponibles en el corpus CAELE/2.
- Analizar los errores desde la perspectiva de AC y AE.

3. Estado del arte

Este apartado presenta una búsqueda de los antecedentes en análisis contrastivo (AC) y análisis de errores (AE) en países hispanohablantes y no hispanohablantes, con el fin de observar los resultados de los análisis contrastivos y el efecto de la cercanía lingüística y geográfica en la producción de errores. Después de esta exploración inicial se indaga sobre las investigaciones ya hechas en AC y AE en India. Esta exploración resalta los puntos claves de cada investigación descrita, que nos darán una perspectiva de lo que ya se hizo en el campo y lo nuevo que se va a explorar en esta investigación. Se hace tal descripción a partir de las tesis de máster y artículos académicos ya publicados en el campo de ELE, con una fuerte base en el análisis de errores en las muestras escritas. Luego, se elabora el tema de la investigación, el paradigma usado para cada estudio, las características de la población, el corpus preferido con el instrumento de recolección de datos (en este caso muestras escritas), el instrumento de análisis y en último lugar las conclusiones de esta exploración. Es importante destacar que la mayoría de las investigaciones en análisis de errores, en general, tiene un alcance teórico, descriptivo y exploratorio y las propuestas o secuencias didácticas de las mismas son escasas.

3.1 Análisis contrastivo en lenguas geográfica y lingüísticamente cercanas al español

Bru (2016) hizo un análisis de calcos sintácticos y errores en el paradigma verbal en estudiantes bilingües en la universidad Nebrija, Barcelona, con 30 estudiantes divididos en dos grupos, grupo A con 15 estudiantes con lengua materna catalán y grupo B con 15 estudiantes con el castellano como lengua materna. Utilizaron un cuestionario para conocer el perfil lingüístico del estudiante. Vale señalar que la taxonomía fue adaptada de las clasificaciones propuestas por Fernández (1997), Vásquez (1999), Atienza, Battaner, Bel et al (1998, 2000), Casanovas (1998) y Balbino de Amorim (2007). En los

resultados, el grupo A tenía muchos errores de calcos sintácticos de su lengua materna catalán. Así, un 53% correspondía a los errores morfosintácticos en las categorías verbales (19%) y calcos sintácticos (21%). De los 122 errores interlinguales identificados, la concentración de errores en unidades aisladas arroja un 62% de errores locales (estudio pragmático). En particular, esta investigación hizo un análisis discursivo a través de las categorías global y local, informando sobre la mediación del mensaje y la posible interrupción de ella y bajo qué criterio.

Torijano (2014) hizo un análisis de errores de estudiantes lusófonos sobre los verbos españoles (I): presente y futuro en la universidad de Salamanca, España. Recogió el corpus de aprendientes de español brasileños. No mencionó la edad de los estudiantes, pero todos tenían el portugués como lengua materna y un nivel de español avanzado. Eligió los aprendientes de los 3 niveles de DELE y Diploma Básico de español (DBE). Ellos tenían que escribir una carta personal y una redacción de las opiniones generales, así como responder a preguntas de comprensión global (850 muestras escritas). En los resultados, mostró que la interlengua tenía componentes culturales de países hispanohablantes que tienen fronteras geográficas con Brasil. Por ejemplo, los vulgarismos y los errores que cometen los nativos hispanohablantes. La taxonomía estaba compuesta por los calcos morfológicos, falsa discordancia, hipergeneralización, diptongación y vulgarización. Este autor también argumentó la lógica de la metodología propuesta por Fernández (1997) y cambió su secuencia. El conflicto más grande era en la morfología de verbos irregulares (especialmente en el pretérito indefinido), vulgarismos, régimen preposicional, solecismos semánticos, diptongación, neutralización errónea del pretérito indefinido y pretérito imperfecto. Su análisis demuestra la interferencia de la lengua materna (en este caso portugués), aun más cuando la L1 y L2 sean de la misma familia lingüística.

Martínez (2016) hizo un análisis de errores en la producción escrita de francófonos aprendices de español como L2 en la Universidad Veracruzana, México. El corpus estaba compuesto por 44 muestras de estudiantes con segundas lenguas diferentes como el francés, el italiano, el portugués, el inglés y otras. Tenían que escribir sobre un tema abierto como - ¿Cómo es y cómo ha sido tu relación con el español? A continuación, se dividió el corpus en 4 categorías, grupo A (16 muestras) tenía inglés y español, grupo B (8 muestras)- inglés, español e italiano, grupo C (10 muestras) inglés, español y portugués y grupo D (10 muestras) - inglés, español y otra. En total hubo 384 errores en tres categorías- léxicos, morfológicos y sintácticos. Dentro de los errores léxicos incluyeron los errores generales, léxicos semánticos y léxicos sintácticos. Los errores léxico sintácticos estaban aun más clasificados a nivel general (80 errores), pronombres (13 errores) y artículos (16 errores). La clasificación de los errores morfológicos tenía dos categorías- morfo ortográficos (general (64), alomorfos (27)), morfo sintácticos (perífrasis (12), concordancia- nominal (28) y verbal (36)), de modo (4) y de tiempo (19). Dentro de los errores sintácticos se incluyeron errores generales (15), de preposición a, en (32), infinitivo (22), objeto directo (7) y complemento de régimen preposicional (4). El autor mencionó cómo la hermandad entre idiomas confunde los matices semánticos de dos términos aparentemente equivalentes, dado que los estudiantes suelen copiar la ortografía de su primera lengua en la segunda o mezcla las sintaxis de ambas.

Paniagua (2005) hizo un análisis de errores e interlengua de aprendices árabes de español como lengua extranjera en la Universidad de Salamanca, España. Eligió las pruebas de los estudiantes con un promedio entre 11.5- 13.5. Al final, eran 60 muestras de 30 estudiantes, 2 muestras por participante. El corpus estaba conformado por las pruebas de los exámenes D.B.E (Diploma Básico de Español). Los participantes redactaron dos cartas: formal e informal de unas 150-200 palabras. En total hubo 2.067 errores, el grupo

libanés con 994 errores y argelino con 1.073 errores. Hubo 139 desviaciones que eran transferencias de la lengua materna; 22 casos eran de omisiones de preposición en las construcciones con complemento directo. 304 desviaciones fueron registradas por transferencias de L2 francés (14,7%) del total de corpus. Dentro de las desviaciones sintácticas que además era la mayoría (125 desviaciones) 63 casos eran de preposiciones, 21 casos de tiempo verbales de francés, 4 casos de pronombres sujeto innecesario y 4 casos de adición de cláusulas relativas innecesarias. Para concluir, 3 de cada cinco errores cometidos tuvieron un origen intralingüístico.

En las cuatro investigaciones: Bru (2016), Torijano (2014), Martínez (2016) y Paniagua (2005), se observó que los autores hicieron un análisis de la población con lenguas maternas cercanas geográfica y lingüísticamente al español. Es de destacar que esta metodología de la delimitación de lenguas maternas en la población meta ayuda en el análisis contrastivo a discurrir sobre el error. Aunque la limitación común en todos los estudios fue la falta de consideración de las otras lenguas que hablaban los estudiantes ya que las poblaciones de los cuatro estudios eran bilingües y/o plurilingües. En los resultados se observa que las lenguas cercanas al español, en este caso: catalán, portugués, francés y árabe, tienen una influencia en el aprendizaje de español en el nivel gramatical, el nivel sintáctico (influencia lingüística) y el pragmático (influencia geográfica). Se pudo observar que muchos de los errores en niveles básicos eran interlingüísticos y con el avance en el nivel de español, empezaron a surgir más errores intralingüales. Esto se puede explicar por la menor dependencia de la lengua materna con el avance en los niveles de la lengua meta.

Es de aclarar que el mito de que las lenguas cercanas a la lengua materna del estudiante son fáciles de aprender resulta falso. Esta afirmación se considera falsa porque los estudiantes, como vimos en los resultados de los cuatro estudios, se confunden entre los componentes cercanos de las dos lenguas: la lengua materna y la lengua meta. De la misma

forma, el estudiante puede enfrentar otros problemas por razones personales y otros factores subjetivos (como la motivación) a la hora de aprender el léxico, la gramática, la sintaxis entre otras competencias lingüísticas, sociolingüísticas y pragmáticas. Por ende, estas investigaciones dan luces sobre la importancia de analizar las lenguas maternas de los aprendientes para entender mejor las necesidades de los estudiantes.

3.2 Análisis contrastivo en lenguas geográfica y lingüísticamente lejanas al español

López (2012) realizó el análisis de errores de interlengua de español para serbio hablantes en las dificultades de uso del artículo. Eligió 15 estudiantes serbio hablantes que tenían entre 22-23 años y hablaban serbio como lengua materna y como L2 y L3 francés e inglés respectivamente. Tenía nivel B1.2 y estaban cursando el cuarto año de Filología francesa de la Universidad de Novi Sad, Serbia. En el corpus se utilizaron las composiciones libres hechas durante el curso, actividades de llenar huecos y se aplicaron entrevistas oportunas para la verificación y complementación de los resultados. Después de hacer un análisis contrastivo entre serbio y español, se elaboró una clasificación de los errores habituales cometidos por los alumnos en el uso de los artículos. En este análisis contrastivo, el autor mencionó los errores de concordancia del determinante (12%), omisión del artículo (36%), omisión del indeterminado (7%), adición del artículo (24%), adición del indeterminado (13%), elección errónea de artículo (5%) y elección errónea de indeterminado (4%). Los resultados demuestran la interferencia del serbio en la cual la categoría de los artículos no existe y el sistema de posicionamiento de los componentes lingüísticos es más flexible.

Castanedo (2019) hizo un análisis de errores durante la adquisición del español en ruso hablantes de nivel B1 en el aula de ELE: el uso del artículo, en la Universidad de Cantabria, España. Se eligieron 30 estudiantes ruso hablantes de nivel B1 pertenecientes a

la universidad de Rostov del Don, 4 estudiantes de máster (teoría de la traducción y comunicación intercultural) y 24 estudiantes de grado (tercer curso de filología hispánica) todos acreditando B1 por el Instituto Cervantes. Eran 32 mujeres y 3 hombres con edades entre 18 y 23 años. También hablaban el inglés y el francés como lenguas extranjeras. Llenaron dos encuestas- reflexión sobre las dificultades en el estudio de la gramática (artículos, tiempos verbales, verbos ser y estar) y ejercicios gramaticales. Los estudiantes poseían problemas con los tiempos verbales, el artículo (el ruso carece de artículos) y los verbos ser/estar (dado que en ruso su uso es diferente). Para ejemplificar, según los resultados, usaron más el artículo determinado que el artículo indeterminado y relacionaron el género de la palabra con el correspondiente al de su lengua materna. También, hubo fallos en el uso del artículo determinado por el indeterminado y viceversa: 30, ausencia del artículo si no era necesario: 28, uso del artículo en aquellos casos en que no era necesario: 12 errores. 22. 77% de los estudiantes dijeron que tenían dificultades en el estudio de la gramática, 22% dijeron que no. Los artículos eran el elemento más difícil, con un 67%, 17% verbo ser y estar, 16% los tiempos verbales. Dentro de las limitaciones del estudio, no se logró un análisis de las diferencias morfosintácticas entre el español y el ruso.

Alexopoulou (2010) analizó errores intralinguales e interlinguales en la interlengua escrita de aprendientes griegos de E/LE en la Universidad Nacional y Kapodistriaca de Atenas. Sobre 1547 errores de 98 pruebas, de los 4 niveles de competencia estudiados, un promedio de 84,7% fueron errores intralinguales y 15,3% fueron interlinguales, de hecho, a partir de nivel B (intermedio), los errores interlinguales disminuyeron y los errores intralinguales, aumentaron. Se explicó que el bajo porcentaje de errores interlinguales fue por las desviaciones morfosintácticas y el punto de partida parecía que no era la lengua materna del aprendiente, sino que dependía de la gramática nuclear guiada por los principios de la Gramática Universal (Corder, 1993).

Liu (2012) estudió los errores y la evolución de la interlengua de los estudiantes sinohablantes. Los participantes eran estudiantes chinos del primer año de la Licenciatura en Lengua Española de la Facultad de Español para no Hispanohablantes (FENHI) de la Universidad de La Habana (UH) y estudiantes ya graduados en 2011 de la misma universidad, 120 (80 mujeres y 40 hombres). Se escogieron muestras de 36 participantes que eran 30% de la matrícula en el curso. Para el desarrollo se utilizaron: histórico-lógico, analítico- sintético, inductivo- deductivo y método empírico: prueba pedagógica. Hubo 579 errores: gráficos (con letra inicial de palabras en mayúscula o minúscula, cambio de letras por confusión de sonidos)-14 errores (2.4%), gramaticales- 466 errores (80.5%), léxicos- 5, 8.6 %), discursivos 49 errores (8.5 %). Los errores interlinguales en la muestra A fueron 346 (36.1%) y 612 intralinguales (63.9%). En la muestra B, 226 fueron errores interlinguales (39.6%) y 344 fueron errores intralinguales (60.4%). Los errores en la muestra B: la omisión del artículo y la concordancia de número, fueron menor que los de la A. Los errores intralinguales predominaron en la producción de textos de ambas muestras. Los errores cometidos con mayor frecuencia por los estudiantes de la muestra A fueron la concordancia de número y de género. Liu mencionó que, “aunque hayan progresado mucho en su estudio del español, la interferencia del chino no se reduce, sino que aumenta por haber terminado el estudio académico de su carrera y no atender con cuidado sus errores en la producción textual.”

Saito (2005) analizó los errores en la expresión escrita en la universidad de Salamanca con estudiantes japoneses, bajo un curso de verano del español para extranjeros (2002). La lengua materna de todos era el japonés y la segunda lengua era el inglés. Tenían que traducir unos textos escritos de japonés a español y así tuvieron 31 muestras escritas al final. En total, tenían 30 minutos para cada pregunta y 4 temas diferentes según el contexto sociocultural de los estudiantes: una carta informal (10 traducciones), un artículo (8

traducciones), una carta formal (4 traducciones) y una receta de cocina (5 traducciones).

Los resultados tenían una tipología: gráficos (130), gramaticales (241), léxicos (97), pragmáticos (19) y la interferencia se dividieron en 3 formas – de japonés, de inglés y de español. Mayoritariamente, los errores surgieron de la lengua japonesa porque hay muchos elementos ausentes en ella que tiene el español, también hubo una interferencia del inglés (gramaticalmente muy capaces en el idioma) porque era su segunda lengua (de la cual aprendieron el alfabeto romano).

Barrera, (2023) investigó los errores en la concordancia nominal del español con un grupo de estudiantes noruegos de ELE en la Universidad de Noruega. Los participantes estudiaron español por 3 años durante la enseñanza secundaria, tenían casi 21 años y hablaban inglés como segunda lengua. Respondieron tres actividades con aspectos diferentes de la concordancia nominal: transparencia morfológica, concordancia adjetival con sintagmas nominales coordinados y la concordancia en posición de atributo. La primera actividad tenía 16 oraciones de concordancia de adjetivo, la segunda también tenía 16 oraciones para estudiar la concordancia, en la tercera, había 18 oraciones para concordar el verbo copulativo y el sintagma nominal formado. En la primera actividad los estudiantes cometieron 154 errores y las respuestas correctas fueron 150. Asimismo, los errores cometidos por la confusión en género fueron 69, errores de número también 69 y errores de terminación 20. En la segunda actividad los errores en total fueron 140 y las respuestas correctas fueron 150. En esta actividad, hubo 54 errores de género, 83 errores de número y 7 errores de terminación. En la tercera actividad, los estudiantes cometieron 185 errores y las respuestas correctas fueron 622, con 20,63% de errores de género, 17,51% errores de número, 1,56% errores de terminación de género y 4,69% errores de redacción. En conclusión, a los estudiantes se les dificultaba el concepto de la concordancia en español, la importancia de sustantivo y la identificación del valor de género de un

sustantivo. Muchas veces los estudiantes tomaban en cuenta solo el género del último sustantivo en la oración sin considerar los otros.

En las seis investigaciones de hablantes del serbio, el ruso, el griego, el chino, el japonés y el noruego, el AC nos ayuda a entender la lejanía lingüística de las lenguas con el español. Muchos de estos idiomas como el ruso, el japonés y el noruego no comparten los mismos componentes gramaticales. A veces hay un elemento más o un elemento menos, por ejemplo, la falta de los artículos que dificulta entender el posicionamiento y su uso. Otro error común fue la concordancia de género y número en el paradigma nominal y la concordancia de número (las conjugaciones) en el paradigma verbal. Estos resultados demuestran cómo la lejanía lingüística y geográfica afecta al aprendizaje de la lengua meta, en este caso el español.

Para concluir, se puede decir que si la lengua meta y la lengua materna de los estudiantes es geográfica y lingüísticamente cercana, el cambio lingüístico es pequeño, asimismo las dos lenguas tendrán más elementos lingüísticos similares y con el mismo valor. Al contrario, si la situación es opuesta y la lengua meta y la lengua materna son lejanas tanto geográfica como lingüísticamente, es muy probable que el sistema lingüístico tenga componentes adicionales, eliminados o con un valor diferente en las dos lenguas. Cabe agregar que el cambio en el sistema lingüístico entre la lengua materna y la lengua meta puede apoyar o irritar a los estudiantes dependiendo de sus estrategias de aprendizaje, las otras lenguas que hablen, entre otros factores subjetivos relacionados con la personalidad de los estudiantes.

3.2.1 Análisis de errores en India

Durante la búsqueda de los antecedentes de análisis de errores, se encontraron pocas investigaciones en la línea de ELE y aun menos que se traten de análisis de errores e interlengua en India. Sin embargo, se encontró una investigación de Rodrigo (2018) sobre

la expresión escrita contextualizada y la corrección guiada en ELE que emplea un análisis de errores implícitamente.

Rodrigo (2018) hizo una investigación con la Universidad de Valencia y Universidad de Mumbai, India con la expresión escrita contextualizada y la corrección guiada en ELE. Tenía 7 estudiantes de grado de francés (20-28 de edad) y una alumna de 65 años, todos de nivel B1. Los participantes tenían que redactar un cuento contextualizado en India, luego, se les pedía corregirse autónomamente desde una reflexión del cuento. La investigación analizó los errores más frecuentes por la interferencia del inglés y también los errores generales. Tenía consideraciones en cuanto a la extensión del cuento, el contexto de la India, los valores de los derechos humanos que podían ser inventados o adaptados y las referencias culturales que hubo que explicarse. En los resultados se mencionó (sin indicar la frecuencia exacta) que los estudiantes cometieron errores de los pretéritos, la ortografía, la ausencia de la preposición, el uso de la preposición incorrecta, la interferencia de inglés (como las mayúsculas), los calcos estructurales, la sintaxis, la confusión de ser y estar, la deixis. En la corrección guiada hubo una dificultad en saber el origen del error y había una ausencia del uso de imperativo. Es de señalar que la investigación solo mencionó el inglés como una lengua común entre los estudiantes, pero no brindó datos de las lenguas maternas que, en general, no suele ser inglés (Bharti, 2019). En resumen, los estudiantes mejoraron las destrezas a lo largo del curso, también se mencionó que hubo una falta de un corpus significativo para los indios y los estudios hechos en hindi hablantes eran escasos.

En India hay estudios en ELE con objetivos diversos y con enfoques diferentes que se han ido cambiando a lo largo de dos décadas. Estos estudios se interesaron inicialmente por la didactización de la literatura: Contextos y libros de texto en la enseñanza de ELE en India (Saxena & González de Lucas, (2004). Luego, se empezó a enfocar más en la didactización de las películas y usar el cine como una herramienta de estudios: *El cine*

como instrumento didáctico en las aulas de ELE en un país de Bollywood (Singh & Mathur, 2010), *El uso de cortometrajes en el aula de ELE: Una propuesta didáctica para el contexto multicultural de la India* (Kalocsányiová, 2012), *Emplazamiento de ciudades en el cine; En la promoción del turismo español: El caso de Zindagi Na Milegi Dobara, Solo se vive una vez, 2011* (Noguero, 2014). Asimismo, las investigaciones más recientes se enfocaron en la situación de ELE en India, los estudios léxicos, el español con fines específicos en turismo y la gamificación: *¿Y tú por qué estudias español? Una propuesta didáctica para el contexto multicultural y multilingüe en el sistema educativo superior de la India* (Álvarez, 2015), *Reflexión en la enseñanza: creencias de seis profesores nativos de ELE en el contexto de India* (Ramallo, 2015), *Préstamos remotos del sánscrito al español: El comercio de piedras preciosas en la India antigua* (Scriba, 2017), *La importancia de la enseñanza de la competencia pragmática en el curso del español con fines específicos para los guías turísticos en la India* (Rehman, 2021), *La gamificación en el aula de ELE* (Kartik, 2021), *Enseñar ELE en la India* (Herrero, 2022).

En cuanto a estas tendencias investigativas, se puede decir que las investigaciones solían partir de la realidad y el uso del español en el país por argumentos lúdicos o laborales. No obstante haber estudios hechos hasta ahora con un enfoque didáctico con varias propuestas didácticas, es evidente la carencia de estudios empíricos en ELE. Dicho esto, todavía se considera que la falta de más estudios en este campo se debe al poco interés de los profesores indios de ELE en temas no literarios y también por la ausencia de los profesores extranjeros que hablen lenguas originarias de India (Rodrigo, 2018). Por lo tanto, esta investigación abre más oportunidades para el análisis de otras perspectivas como el análisis de errores en la población india.

4. Justificación

La justificación de esta investigación se puede dar desde la relevancia social, la implicación práctica y la utilidad metodológica.

Primero, desde la relevancia social, esta investigación podría aportar a la creación de material didáctico para los estudiantes indios. Este análisis de errores tendrá un valor tanto cualitativo como cuantitativo para los docentes de español en la India (aún más para los docentes extranjeros que no saben hindi), luego el planteamiento de currículos también podría beneficiarse de este análisis. La investigación permitirá visibilizar a la población india en los procesos de enseñanza y aprendizaje del español.

Por otra parte, durante la búsqueda en las bases de datos en las investigaciones de análisis de errores, se encontró un vacío investigativo en ELE en la población india. Tampoco hay una recopilación de muestras escritas u orales para realizar estos trabajos. Así que, este proyecto también busca recopilar un corpus de muestras escritas de aprendientes indios de ELE en el corpus CAELE/2 del Instituto Caro y Cuervo. Esta herramienta ayudaría a la visibilización de las necesidades de la población india para las investigaciones en el futuro y luego su posible explotación didáctica.

Segundo, desde la implicación práctica, la investigación aporta a que los profesores se anticipen a las posibles fosilizaciones y confusiones comunes entre los estudiantes a nivel morfosintáctico del español como lengua extranjera. Del mismo modo, los patrones de errores encontrados en el corpus sensibilizarán tanto a los docentes extranjeros como a los docentes originarios de la India, para saber cómo aprovechar los conocimientos lingüísticos ya desarrollados por parte de los estudiantes.

Tercero, desde la utilidad metodológica, los instrumentos desarrollados a lo largo de este proyecto tienen un valor investigativo e impulsan a las posibles investigaciones en

India. De la misma forma, para propuestas a futuro, se puede adaptar la metodología elegida y sustentada en este estudio con los estudiantes indios, dado que, el corpus recopilado con las diferentes lenguas maternas ofrece una oportunidad para explorar y analizar los errores desde otras perspectivas y realizar análisis contrastivos en otros paradigmas.

5. Marco teórico

En este apartado se recopila la teoría y los conceptos de interés para esta investigación. Por lo tanto, se presentan los referentes teóricos pertinentes en este campo: las hipótesis de análisis contrastivo, análisis de errores y la interlengua. En el análisis contrastivo, se explorará el contraste exhaustivo entre hindi y español a nivel morfosintáctico en el paradigma nominal. Luego, se presentará la lingüística de corpus como una metodología y algunas taxonomías para el análisis contrastivo y el análisis de errores pertinentes para esta investigación.

5.1 Análisis contrastivo

En la década de los 50 y 60 se difundieron los postulados conductistas que se enfocaban en la repetición y la práctica para lograr la fluidez en la lengua meta (Fries, 1945). Luego, Lado (1957) propone comparar el sistema lingüístico de la lengua materna del estudiante con la lengua meta que quiere adquirir bajo la teoría de la transferencia de lenguas (*Language transfer*). Sin embargo, Skinner (1972) defendió la postura conductista y estableció que, para adquirir una lengua, el aprendiente tiene que seguir una serie de hábitos basados en el esquema estímulo-respuesta-refuerzo.

Después de esta propuesta inicial de Lado en 1957, surge la hipótesis de Análisis contrastivo (AC), cuya idea está directamente ligada a la distancia interlingüística: la distancia entre la lengua materna y la lengua meta. Así, se afirma que cuanto mayor sea la distancia, mayor será la dificultad en el aprendizaje y mayor también la posibilidad de la interferencia. Por ejemplo, aprender portugués será más fácil para un hispanohablante por la misma cercanía de las dos lenguas, pero aprender coreano para el mismo hispanohablante sería más difícil por la distancia lingüística entre los dos idiomas. Sin embargo, se observó que en la investigación de Torijano (2014) con los lúsofonos, los aprendientes se

confundieron en la morfología de los verbos irregulares (especialmente en el pretérito indefinido), vulgarismos, régimen preposicional, solecismos semánticos, diptongación, neutralización errónea del pretérito indefinido y pretérito imperfecto. En este orden de ideas, se puede decir que el portugués, que es lingüística y geográficamente cercano al español, también tiene sus retos durante el proceso del aprendizaje.

Ahora bien, estas similitudes y diferencias entre la lengua materna y la lengua meta pueden señalar las áreas de posibles dificultades y predecir los problemas que el estudiante va a enfrentar durante el proceso del aprendizaje de un segundo idioma. Por tanto, los resultados de este análisis pueden ayudar a los docentes en la creación del material didáctico y técnicas de instrucción con base en las necesidades de los estudiantes. Fries postula esta misma idea (1945, 9) en la siguiente frase: “Los materiales más eficientes son aquellos basados en una descripción científica de la lengua que vamos a estudiar, cuidadosamente comparada con una descripción de la lengua del estudiante”. Esto afirma que los estudiantes tienden a transferir las estructuras y el vocabulario de su lengua materna a la lengua aprendida. Por eso, se puede proponer un procedimiento para realizar este análisis contrastivo en todos los niveles del sistema lingüístico: fonológico, gramatical y léxico siguiendo el proceso propuesto por Gargallo (1998), que se expone a continuación:

- a. Descripción estructural de L1 y L2;
- b. Cotejo de las descripciones;
- c. Elaboración de un listado preliminar de estructuras no equivalentes;
- d. Reagrupación de las mismas estableciendo una jerarquía de dificultad;
- e. Predicción y descripción de las dificultades;
- f. Preparación de los materiales de instrucción.

En 1957, las predicciones de Lado se rechazaron porque consideraban el error un obstáculo para adquirir la lengua meta; pero, aun así, sus estudios son la base para comprender los actuales estudios en la interlengua. Su objetivo era construir una gramática

contrastiva con el fin de establecer una jerarquía de correspondencias de la gramática (en este análisis morfosintáctico) y guardar las dificultades en el aprendizaje y las posibles interferencias.

Ahora bien, los dos conceptos propuestos por Weinrich (1954): *interferencia* y Haugen (1953): *préstamo lingüístico*, se inspiran de los estudios sobre la integración de los inmigrantes de los Estados Unidos (EE. UU), que también estudian la influencia de la lengua materna en la construcción de la interlengua. El concepto de *interferencia* de Weinrich se refiere a las desviaciones en la lengua meta que tiene normalmente un bilingüe por la familiaridad con otras lenguas conocidas o en contacto. Por otro lado, el concepto de *préstamo lingüístico* de Haugen se refiere a la difusión cultural y posterior integración de las unidades o rasgos lingüísticos de la lengua A, el cual existía en la lengua B y que A no poseía. Por ejemplo, las palabras como “estándar”, “fútbol” etc.

Al analizar las propuestas de Weinrich y Haugen, se observa que estos toman en cuenta no solo el efecto que tiene la lengua materna en el aprendizaje de la lengua meta, sino también problemas que surgen en la lengua misma (no se consideraba esta propuesta en AC). Por ello, se puede hablar de dos tipos de interferencia:

a. Interferencia Interlingüística (Externa)

La lengua materna influye en los componentes lingüísticos de la lengua meta. Por ejemplo: traducciones literales de lengua materna a lengua meta, el orden sintáctico y ausencia o adición de ciertos componentes gramaticales en la oración, etc.

b. Interferencia Intralingüística (Interna)

La lengua meta que se está aprendiendo es la responsable de los errores. De aquí nace el concepto de *Análisis de errores* del que vamos a hablar en otro apartado del marco teórico.

A continuación, se expondrá el análisis contrastivo entre el hindi y el español en el paradigma nominal. Este análisis contrastivo demuestra primero cada categoría del sistema morfosintáctico de hindi, luego de español y al final el contraste entre los dos.

5.1.1 Contraste entre el sistema morfosintáctico del español y el hindi en el paradigma nominal

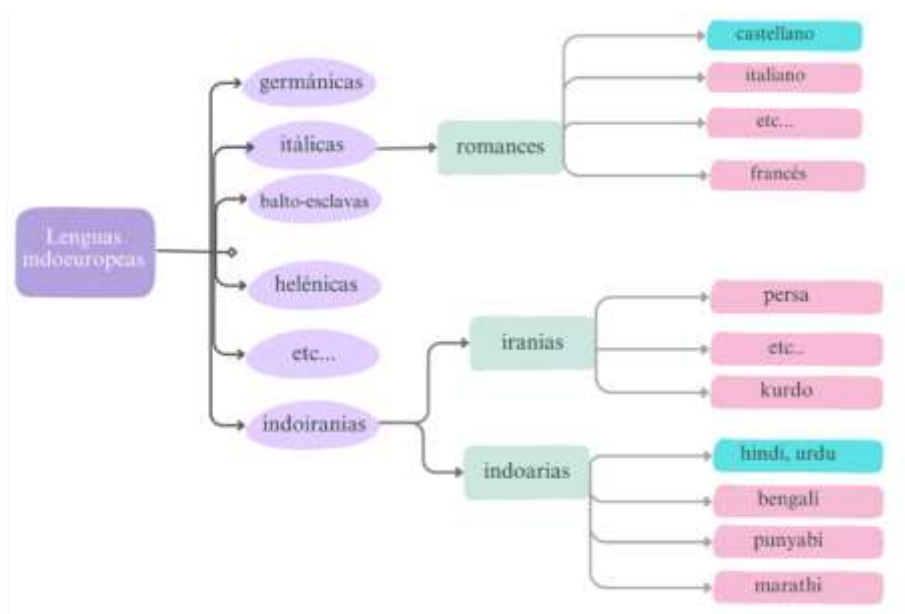


Figura 1: Representación gráfica de la relación entre hindi y español. *Adaptado del libro: Gramática de hindi y urdu para castellano hablantes (Díaz, 2023)*

El hindi y el español pertenecen al grupo de lenguas indoeuropeas. Las lenguas como hindi, bengalí, punyabí y marathi son de la rama de las lenguas indoarias que además pertenecen al grupo de lenguas indoiranias que es la rama más oriental dentro de las lenguas indoeuropeas (Díaz, 2023).

El hindi y el español tienen los mismos tipos de palabra aproximadamente¹: sustantivos, adjetivos, verbos, adverbios, posposiciones (el equivalente a preposiciones en español), conjunciones, verboides, pronombres, partículas diversas (sufijos, prefijos, interjecciones, onomatopeyas) y artículos. En este estudio se expondrán los siguientes componentes

¹ Esta información se ha tomado mayoritariamente del libro –La gramática de hindi y urdu para castellano hablantes escrito por Gonzalo Díaz (2023) y Estructuras gramaticales de hindi y español por Vasant Ganesh Gadre (1996), así como del conocimiento personal del autor.

morfosintácticos: flexión (accidentes gramaticales), sustantivos, adjetivos, artículos, determinantes y orden sintáctico del sintagma nominal.

5.1.1.1 Flexión

La flexión en la lingüística se refiere al cambio de forma de una palabra para expresar ciertas características gramaticales. Asimismo, las propiedades responsables de este cambio en la palabra se llaman *accidentes gramaticales*. Por ejemplo, en español los adjetivos pueden tener dos *accidentes gramaticales* : el género y el número (tabla 1)

Tabla 1: *La flexión de género y número en adjetivos en español*

		Género	
Número	Adjetivo- pequeño	Masculino	Femenino
	singular	pequeño	pequeña
	plural	pequeños	pequeñas

Nota. Elaboración propia

Con respecto al hindi, la flexión es del género, el número y el caso. Existen solo dos géneros en hindi como en español: el masculino y el femenino. También, igual que en español, en hindi existen dos números: singular y plural. Sin embargo, se ha perdido en español la tercera categoría, el caso: son los cambios que se efectúan en un sintagma dependiendo de la función sintáctica de esa palabra en una oración. El latín (la lengua madre del español) tiene seis casos y sánscrito (la lengua madre del hindi) tiene ocho casos de los cuales el hindi conserva los siguientes tres:

- a. Directo: El caso *no marcado* o *nominativo*. Usado, en general, para palabras que funcionan como el núcleo del sujeto.

- b. Oblicuo/ indirecto²: Es la palabra o grupo de palabras (sintagma) la/el que desempeña una función como una palabra o construcción que está subordinada al resto de la oración mediante nexos (posposiciones). También se usa en ciertas ocasiones cuando la palabra funciona como complemento circunstancial en el predicado.
- c. Vocativo: Es la palabra o grupo de palabras (sintagma) la/el que desempeña que representa exclusivamente inserciones vocativos dentro de una oración, para darle énfasis en el discurso, dirigido a la segunda persona.

Las palabras que experimentan flexión de caso suelen flexionarse también de acuerdo con el género y/o número y este fenómeno de los cambios de una palabra por flexión de caso, número y género se llama *declinación*. En hindi, la mayoría de los tipos de palabra se declina. A continuación, en las siguientes categorías solo se va a explicar los dos casos: directo e indirecto porque el vocativo tiene una función discursiva y no morfosintáctica.

5.1.1.2 Sustantivos

Español

En el español los sustantivos tienen flexión de género y número. El sistema de expresión del número en los sustantivos españoles es más regular que el del género y está formado por la simple oposición del singular con el plural. El plural se marca por los alomorfos -s, -es, frente al singular que se marca con Ø. (Marticorena (2011)). Se puede ver este sistema de número en la tabla 2.

Tabla 2: El sistema de números en los sustantivos

Singular	Plural
Pájaro	Pájaros -s

² diferentes autores eligen entre el primero o el segundo término. En esta investigación se usa la palabra indirecto.

Balcón	Balcones -es
Sintaxis	Sintaxis - Ø

Nota. Tomado y adaptado desde el libro: Estructuras gramaticales de hindi y español (Gadre, 1996)

En cuanto al género en español, todos los sustantivos se pueden categorizar en masculino o femenino, aunque no todos tienen la flexión de género. Generalmente, los sustantivos que terminan en –o son masculinos y los que terminan en –a son femeninos. Pero no siempre funciona esta regla, además, no todos los sustantivos terminan en –o y –a. En todo caso, en el idioma español, no existe un género neutro. Conviene advertir que no hay formas de expresión fijas del género comunes a todos los sustantivos. Suele decirse que estas son las reglas:

-a/-d/-z: femenino

-a/-o/-i/-u/-consonante: masculino.

Por su parte, la Real Academia Española enumeró varias reglas para determinar el género de los nombres por sus terminaciones y cada regla tiene sus excepciones. En cualquier caso, la concordancia con un adjetivo adyacente (que tenga flexión genérica) y con el artículo (salvo a los nombres que contenga flexión genérica) es el auténtico sistema para reconocer el género (Quilis y Hernández (1978)). En este orden de ideas, el sustantivo ejerce la subordinación en el sintagma nominal y los artículos y los adjetivos concuerdan con él según su género. Sin embargo, el género del sustantivo solo (no en una oración), suele resultar difícil de predecir por la falta de reglas generalizadas. Por lo tanto, no existen reglas fijas y específicas para la asignación de género.

Hindi

Los sustantivos en hindi se flexionan o declinan según su número y su caso. Los casos tienen diferentes nombres según los autores. En el caso directo, los sustantivos no van

acompañados por posposiciones y tienen la función, generalmente, de sujeto u objeto directo. Pero, en el caso indirecto, el sustantivo es seguido por una postposición. Las funciones expresadas por los casos son varias como: sujeto, objeto, modo, propósito, ventaja, separación, origen, posesión, material, composición, lugar, tiempo, etc.

En el caso indirecto, la forma se modifica indicando las relaciones con los demás componentes, que se expresan a través de diferentes posposiciones como: *ne*, *ke*, *se*, *men*, *par* y *kaa*, que también se llaman marcadores de caso o *posposiciones*. Por ejemplo: *ladke ko* (al chico), derivado de *ladkā* (chico), *gharon men* (en las casas) derivado de *ghar* (casa). Abajo se muestran algunos ejemplos tomados de Díaz (2023) para ilustrar estas diferencias.

Declinaciones típicas de los sustantivos en hindi

Los sustantivos que terminan en –a como *ladkā* (muchacho), son típicamente masculinos.

Tabla 3: Declinaciones de *ladkā* (muchacho)

Ladkaa (muchacho)	singular	plural
directo	<i>ladkā</i> / लड़का	<i>ladkē</i> / लड़के
indirecto	<i>ladkē</i> / लड़के	<i>ladkōm</i> / लड़कों

Nota. Adaptado desde el libro: Gramática de hindi y urdu para castellanohablantes (Díaz, 2023)

Los sustantivos que terminan en –i como en *bēṭī* (hija) son típicamente femeninos.

Tabla 4: Declinaciones de *beti* (hija)

Beti (hija)	singular	plural
directo	<i>bēṭī</i> / बेटी	<i>bēṭiyāṁ</i> / बेटियाँ
indirecto	<i>bēṭī</i> / बेटी	<i>bēṭiyōn</i> / बेटियों

Nota. Adaptado desde el libro: Gramática de hindi y urdu para castellanohablantes (Díaz, 2023)

Muchos otros sustantivos no tienen ninguna de estas dos terminaciones típicas y el género no se puede deducir de la terminación. Por ejemplo, *gēēt* (canción) es masculino, pero *bāt* (palabra) es femenino.

Tabla 5: *Declinaciones de geet (canción)*

Geet (canción)	singular	plural
directo	<i>gēēt</i> / गीत	<i>gēēt</i> / गीत
indirecto	<i>gēēt</i> / गीत	<i>gēētōn</i> / गीतों

Nota. Adaptado desde el libro: Gramática de hindi y urdu para castellanohablantes (Díaz, 2023)

Tabla 6: *Declinaciones de baat (palabra)*

Baat (palabra)	singular	plural
directo	<i>bāt</i> / बात	<i>Bātēin</i> / बातें
indirecto	<i>bāt</i> / बात	<i>Bātōn</i> / बातों

Nota. Adaptado desde el libro: Gramática de hindi y urdu para castellanohablantes (Díaz, 2023)

Para concluir, se puede decir que en hindi existen cuatro formas de reconocer el género de los sustantivos dependiendo de la terminación.

1. Masculinos con una terminación – ā. Por ejemplo- *ladkā* (muchacho)
2. Femeninos con una terminación –i. Por ejemplo- *bēti* (hija)
3. Masculinos cuyo género hay que recordar (no terminan en –a). Por ejemplo: *gēēt* (canción)
4. Femeninos cuyo género hay que recordar (no terminan en –i). Por ejemplo: *bāt* (palabra)

De todas las declinaciones se pueden extraer algunas reglas prácticas.

1. Las terminaciones de los plurales de todos los indirectos son siempre –on.

2. Todas las formas del singular son iguales, excepto los sustantivos con género masculino.

Declinaciones atípicas de los sustantivos en hindi

Hay pocos sustantivos, pero con alta frecuencia de uso, que terminan en –aa o en –i, pero pertenecen al género opuesto. Por ejemplo – *mātā* (madre) es femenino y *ādmī* (hombre) es masculino. Estos sustantivos no varían sus terminaciones en el singular, pero por lo demás se declinan.

Tabla 7: Declinaciones de *ādmī* (hombre)

Aadmi (hombre)	singular	plural
directo	<i>ādmī</i> / आदमी	<i>ādmī</i> / आदमी
indirecto	<i>ādmī</i> / आदमी	<i>ādmīyōn</i> / आदमियों

Nota. Adaptado desde el libro: Gramática de hindi y urdu para castellanohablantes (Díaz, 2023)

Tabla 8: Declinación de *maataa* (madre)

Maataa (madre)	singular	plural
directo	<i>mātā</i> / माता	<i>mātayēin</i> / माताएँ
indirecto	<i>mātā</i> / माता	<i>mātayēin</i> / माताओं

Nota. Adaptado desde el libro: Gramática de hindi y urdu para castellanohablantes (Díaz, 2023)

Comparación entre hindi y español

En comparación con el español, las reglas en hindi tienen ciertas características diferentes por la involucración de las declinaciones y sus combinaciones. No obstante, los estudiantes tienen que memorizar estas reglas y sus excepciones en las dos lenguas por la falta de una regla general sin excepciones. En resumen, las características generales de estos patrones atípicos en hindi son:

1. Todas las formas del singular son iguales, sin importar el caso.

2. En accidente gramatical de palabras masculinas, el plural directo también es igual que los singulares.

Los conceptos del número y del género existen en las dos lenguas. Eso quiere decir que los estudiantes hindi hablantes ya están familiarizados con estos conceptos antes de empezar a aprender español. También saben que la formación de plural se hace, en general, agregando un morfema en la terminación de la palabra singular (que no pasa en el caso de inglés). Por lo tanto, esto puede ayudar a reducir interferencias en el aprendizaje de español por parte de los hindi hablantes. Dado que, tanto en sánscrito como en latín, hay tres géneros: masculino, femenino y neutro y la forma aleatoria de la asignación del género a los sustantivos neutros en latín y sánscrito en español e hindi respectivamente no tiene explicación en ninguno de los dos idiomas. A pesar de una distinción morfosintácticamente clara a partir de la flexión en los morfemas, hay interferencias que, al no ser de tipo semántico o morfosintáctico, suelen ser de tipo morfológico porque se puede observar que algunas palabras masculinas en español son femeninas en hindi y viceversa (figura 3).

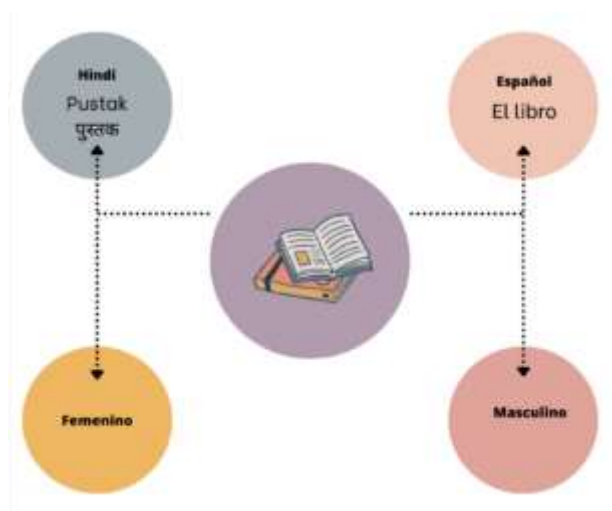


Figura 2: Contraste del género del sustantivo libro en español e hindi, ilustración propia.

5.1.1.3 Adjetivos

Español

Los adjetivos en español tienen flexión en las dos categorías como en el caso del sustantivo: el número y el género, aunque como los sustantivos, no todos se flexionan, por ejemplo: gratis, verde etc. En el caso de la flexión de número el adjetivo concuerda con el sustantivo sin tener muchas excepciones (Gadre, 1996). En cambio, en cuanto a la flexión de género (si el adjetivo se flexiona) se necesita una modificación entre los sufijos de género -o y -a que aparecen al final de la raíz. Por ejemplo (tabla 9):

Tabla 9: Flexión de género y número en el adjetivo 'blanco' y 'bueno'

Singular		Plural	
Masculino	Femenino	Masculino	Femenino
blanco	blanca	blancos	blancas
bueno	buenas	buenos	buenas

Nota. Ilustración propia

Sin embargo, como se ha aclarado al principio, hay muchos adjetivos que no tienen flexión de género, solo de número y pueden terminar en las vocales *a, e, i* o en una de las consonantes *n, l, r, o, z*; como es el caso en *dulce, común, azul, agrícola, capaz* etc.

Hindi

En hindi, los adjetivos se declinan según su género, su número y su caso (Díaz, 2023). De acuerdo con este autor, los adjetivos se pueden dividir en dos grandes grupos:

- La forma paradigmática termina en -a
- La forma paradigmática no termina en -a

Los adjetivos cuya forma paradigmática termina en -a, se flexionan como se menciona en los siguientes puntos:

- Usan la forma en -a para el masculino directo singular.

- Tienen una forma en –e que se usa para todas las otras formas masculinas (también los plurales, sin importar su caso o número).
- Tienen una forma en –i que se usa para todas las formas femeninas (sin importar su caso o número).

Tabla 10: *Declinación de lamba (alto, largo): cuya forma termina en -a*

Lambaa (alto, largo)	caso	masculino	femenino
लम्बा			
singular	directo	लम्बा / lamba	
	indirecto		लम्बी/ lambi
plural	directo	लम्बे/ lambe	
	indirecto		

Nota. Tomado desde Díaz (2023)

Los adjetivos cuya forma paradigmática no termina en –a, se flexionan como se menciona en los siguientes puntos:

- Los de singular terminan igual sin importar el género y el caso.
- Los de plural con el caso directo terminan en –en (nasalizado) sin importar el género.
- Los de plural con el caso indirecto terminan en –on sin importar el género.

Tabla 11: *Declinación de mushkil (difícil): cuya forma no termina en -a*

Mushkil (difícil)	caso	masculino	femenino
मुश्किल			
singular	directo		

	indirecto	mushkil/ मुश्किल
plural	directo	mushkilēn / मुश्किलें
	indirecto	mushkilōn / मुश्किलों

Nota. Tomado desde Díaz (2023)

Comparación entre hindi y español

En los dos idiomas hay flexión de género y número. No obstante, en hindi la flexión está dada también desde el caso. Es evidente desde las tablas 12 y 13 que la flexión en el género en los adjetivos en hindi existe, pero con baja presencia y frecuencia de uso en comparación con la flexión del número o el caso. En este orden de ideas, un estudiante con lengua materna hindi puede tener más dificultades en cuanto al sistema de género y la concordancia con el sustantivo en español.

5.1.1.4 Artículos

Español

Los artículos en español se clasifican en indefinidos y definidos, también conocidos como indeterminados y determinados respectivamente. Sin embargo, se puede agregar una categoría más de *omisión de artículo* para explicar las situaciones y razones por las que, en algunos momentos, el sustantivo no requiere ninguna de las dos anteriores (García, 2005).

1. Definidos (el, la, los, las)

Tabla 12: *Artículos definidos en español*

	<i>Masculino</i>	<i>Femenino</i>
<i>Singular</i>	el	la
<i>Plural</i>	los	las

Se usa el artículo definido para indicar algo que es identificable para el oyente entre todos los otros objetos, por ejemplo³:

- *Este es Pedro, **el** hermano de María.*
(El contexto implícito es que Pedro es el único hermano de María.)
- *No cabemos en mi carro, ustedes tres, **los** amigos de Ana y yo, somos muchos.*
(El contexto implica que los amigos de Ana ya eran conocidos.)

Otro uso de los artículos definidos es de generalizar una clase de objetos. Usamos el, la, los, las.

- *El perro es **el** mejor amigo del hombre.*
(El perro y el hombre representan sus grupos y una relación general entre ellos.)
- *No me gustan **los** gatos, me dan miedo.*
(Los gatos, en general)

2. Indefinidos (un, una, unos, unas)

Tabla 13: Artículos indefinidos en español

	Masculino	Femenino
<i>Singular</i>	un	una
<i>Plural</i>	unos	unas

Se usa el artículo indefinido para indicar algo que no es identificable para el oyente entre otros objetos de su clase, es decir, cuando hay varios objetos del mismo tipo o cuando hablamos por primera vez de ese objeto, por ejemplo:

- *Este es Pedro, **un** hermano de María.*
(El contexto implica que María tiene más hermanos que Pedro.)
- *También vienen **unos** amigos de Ana: Pedro y Juan.*

³Para mayor información se puede referir al libro: Gramática básica de español (A1-B2)

(El artículo tiene una función partitiva, son unos amigos dentro de un grupo de amigos.)

Otro uso de los artículos indefinidos es de indicar el objeto representativo de su clase.

Usamos solo *unos*, *unas* para los sustantivos que siempre existen en forma plural (tijeras, gafas, etc.)

- *El hombre es **un** mamífero.*

(Se habla de un individuo cualquiera que representa una clase, que en este caso es mamífero..)

- ***Un** médico tiene que conocer bien el cuerpo humano.*

(Se habla en general de un individuo cualquiera que representa una clase. Es decir, cualquier individuo miembro de la clase “médico” tiene que conocer bien el cuerpo humano.)

3. Omisión de artículo

Se omite el uso de los artículos en español cuando no nos referimos a ningún objeto concreto. Tampoco los usamos en el caso de una cantidad indeterminada que no necesitamos especificar.

- *Mi casa está bastante cerca. Son diez minutos en Ø **bicicleta**.*
(No hablan de un objeto particular, sino de un medio de transporte.)
- *¿Venden ustedes Ø **alfombras**?*
(No hablan de alfombras de alguna marca o material específicamente.)
- *Si vas al súper, compra Ø **aceite** y Ø **café**.*
(No hablan de aceite o café particular, la persona puede comprar cualquier marca y cualquier cantidad de estos productos.)

Alomorfos

Cuando un sustantivo femenino empieza con “a” o “ha” tónicas, se usa el artículo determinado masculino “el”, como es el caso de *el agua*. Sin embargo, esta regla solo se aplica cuando el sustantivo está inmediatamente después del artículo. Cuando hay otra palabra que lo separa como un adjetivo, se usa el artículo en femenino, por ejemplo: *el*

agua pero la mejor agua. Además, como el sustantivo es femenino, el adjetivo pospuesto también concuerda en el género, por ejemplo: *el agua pura*, *el azúcar morena*, entre otros ejemplos.

Hindi

En hindi, menciona el autor Gadre (1996), no tiene artículos de ninguna clase. Sin embargo, hay casos de traducciones en donde se encuentra el artículo indefinido representado por *ek* (uno). Es decir, la forma frecuente de expresar el artículo en hindi es a través del número *uno* que, en general, no se traduce y no es un elemento esencial. Por ejemplo:

- *Se acerca un caballo.* (Se menciona el número y también artículo en la traducción)
- एक घोड़ा पास आ रहा है।
- *Ēk ghōdā pās ā rahā hai.*
- El caballo cerca viniendo está.

En el siguiente ejemplo, los dos casos son válidos en español: con artículo o sin artículo, pero en el caso de hindi es sin artículo.

- *Es difícil vivir en Delhi sin tener un carro.* (con artículo en español)
- *Es difícil vivir en Delhi sin tener carro.* (sin artículo en español)
- कार के बिना दिल्ली में रहना मुश्किल है।
- *Kār kē binā dillī mēn rēhnā mushkil hai.* (sin artículo en hindi)
- Coche sin Delhi en vivir difícil es.

Dado que no existe una academia de hindi como la Real Academia Española para el español, el mismo autor Gadre (1996) hizo una encuesta para determinar la posición de *ek* (uno) en las oraciones. La realizaron en una clase de Lingüística Aplicada, donde presentaron 4 oraciones. La numeración era solo para las oraciones gramaticalmente adecuadas en las dos lenguas y no para las traducciones literales⁴:

⁴ La primera oración (numerada) es en hindi con letras latinas y al lado en devanagari, la segunda oración es la traducción literal con el orden sintáctico similar al hindi y la tercera oración (numerada) es la oración como debe ser en español.

1. *Main kal ēk film dēkhnē gayā thā.* (मैं कल एक फिल्म देखने गया था।)

Yo ayer una película ver ido había.

2. (Había ido a ver una película ayer.)
3. *Main kal film dēkhnē gayā thā.* (मैं कल फिल्म देखने गया था।)

Yo ayer película ver ido había.

4. Había ido a ver una película ayer.
5. *Āp sē milnē ēk sajjan āyē hain.* (आप से मिलने एक सज्जन आये हैं।)

Vd. a ver un señor venido ha.

6. Ha venido un señor a verle.
7. *Āp sē milnē kōī sajjan āyē hain.* (आप से मिलने कोई सज्जन आये हैं।)

Vd. a ver algún señor venido ha.

8. Ha venido un señor a verle.

Mencionó el autor Gadre (1996) que en las oraciones 1 y 3, el sentido semántico es similar pero, hay cierta variación sintáctica. La oración 1 contiene *ek* (uno), pero la 3 no la tiene. Lo mismo sucede en las oraciones 5 y 7, el 5 contiene *ek* (uno) pero el 7 tiene un determinante indefinido. Los 26 encuestados tenían que responder si las dos variantes en ambos casos eran lingüísticamente aceptables y, en el caso afirmativo, qué variantes preferirían usar en su discurso como equivalentes de las oraciones inglesas:

- *I had gone to see a film yesterday.*

(Había ido a ver una película ayer.)

- *A gentleman is here (has come) to see you.*

(Ha venido un señor a verle.)

En los resultados de la encuesta, los 26 participantes consideraron que las cuatro oraciones eran aceptables lingüísticamente. 22 estudiantes consideraron que la oración 3 era más aceptable entre 1 y 3, es decir, la oración sin *ek* (uno). Solo 4 dijeron que la 1 era más usual y preferían usarla. En la otra pareja de oraciones (5 y 7), ocho estudiantes dijeron que preferían la 7 porque tenía *kōī* (el determinante indefinido), mientras que cuatro se mostraron a favor de la 5 por tener *ek*. El resultado mostró que el uso de *ek* en las oraciones en hindi es una interferencia de inglés que no se usa propiamente en lengua hindi.

En el mismo orden de ideas, en el hindi no hay artículos definidos. Respecto a este punto, dice Tiwari (1979) que, aunque no haya artículo determinado, se introduce el elemento definido cambiando el orden sintáctico. Por ejemplo⁵:

1. *Tumnē pattiyān tōdīn* (तुमने पत्तियाँ तोड़ीं।) Tú hojas cortaste.

2. Tú cortaste unas hojas.

3. *pattiyān tumnē tōdīn*. (पत्तियाँ तुमने तोड़ीं।) Hojas tú cortaste.

4. Tú cortaste las hojas.

5. *rām nē sānp mārā*. (राम ने सांप मारा।) Ram serpiente mató.

6. Raam mató una serpiente.

7. *Sānp rām nē mārā*. (सांप राम ने मारा।) Serpiente Ram mató.

8. Ram mató la serpiente.

9. *Mēz par kitāb hai*. (मेज़ पर किताब है।) Mesa en libro está.

⁵ Se numera solo a las oraciones gramáticas y no las traducciones literales.

10. Hay un libro en la mesa.

11. *kitāb mēz par hai*. (किताब मेज़ पर है।) Libro mesa en está.

12. El libro está en la mesa.

Los ejemplos del 1 al 8 parecen tener la justificación previamente dicha, pero las oraciones 9 y 11 parecen pertenecer a la misma categoría de artículos, es decir, la palabra mesa (*mēz*) sigue teniendo artículo definido sin importar su posición en la oración. Esto sugiere que el razonamiento de Tiwari (1979) solo puede aplicarse al sujeto de la oración. Con base en todos estos ejemplos, se puede decir que no hay artículos definidos en hindi y la posición inicial del sustantivo, especialmente si es el sujeto de la oración, añade un valor definido al sustantivo (Gadre, 1996).

Comparación entre hindi y español

Tabla 14: *Artículos en hindi y español*

Hindi	Español
	el, la, los, las
	un, una, unos, unas
∅	∅

Nota. Adaptado de Gadre (1996)

En resumen, se puede decir que en algunos casos hay una correspondencia de *ek* (uno) de hindi a *un* (artículo indefinido) de español. Ahora bien, como en hindi, el concepto de artículos no existe morfológicamente como en español, hay una alta posibilidad de confusión y dificultades por parte de los estudiantes hindi hablantes que llevan estas interferencias a nivel morfosintáctico durante el aprendizaje de español.

5.1.1.5 Determinantes

Demostrativos

Tabla 15: Los demostrativos en hindi y español

Español				Hindi					
				Caso directo		Caso indirecto		Caso indirecto	
Singular		Plural		Singular	Plural	Singular	Plural	Singular	Plural
Masc.	Fem.	Masc.	Fem.	Mas. /Fem.	Mas. /Fem.	Mas. /Fem.	Mas. /Fem.	Mas. /Fem.	Mas. /Fem.
Este	Esta	Estos	Estas	yah	ye	is	inhōn	is	in
Ese	Esa	Esos	Esas	wah	we	us	unhōn	us	un
Aquel	Aquella	Aquellos	Aquel las						

Nota. Tomado y adaptado desde Gadre (1996)

Español

Según García (2005), en español los demostrativos pueden ser femeninos, masculinos y neutros. Las formas de masculino y femenino concuerdan en género y número con el sustantivo al que se refiere. Los demostrativos sirven para señalar cosas en relación con tres espacios diferentes, entre los cuales todas las flexiones de *este* representan el espacio “aquí”, *ese* representa el “ahí” y *aquel* representa el “allí”. También los usamos para referirnos a cosas que se sitúan en el espacio de la persona con la que hablamos. Por ejemplo:

- *Este bolígrafo no funciona, ¿me dejas ese rotulador?*

Ya se sabe que los demostrativos sirven para ubicarnos o ubicar algo en el espacio, pero también sirven para situar algo en relación con el tiempo. Por ejemplo:

- *Ese día estaba muy cansada, por eso no te llamé.* (El día pasado del que hemos hablado.)

Hindi

En hindi, solo hay dos espacios (*aquí* y *allí*) en vez de tres como en español. Se pueden comparar estos demostrativos con los de inglés (*this* y *that*) porque no representan un espacio intermedio, sino solo dos extremos, en donde los dos son espacios subjetivos que dependen de la perspectiva del hablante. En el caso directo son: *yah* y *wah* y sus formas plurales: *ye* y *we*. En el caso indirecto, los dos demostrativos son: *is* y *us* y sus formas plurales en el ergativo⁶ son: *inhōne* y *unhōne* y en los demás casos indirectos: *in* y *un*.

Comparación entre hindi y español

Se puede observar que la flexión genérica y numérica de los demostrativos en hindi no es tan uniforme como la del español. Aunque los hindi parlantes conozcan un sistema más complejo con declinaciones en el caso, los estudiantes podrían tomar tiempo en entender los tres espacios en español. Dado que es una nueva forma de ver la realidad y necesita un cambio de perspectiva y su respectiva consolidación. Otra dificultad podría ser la concordancia de número y género como se había visto en el caso de los adjetivos.

Posesivos

Español

⁶ Cuando funciona como el agente de un verbo transitivo conjugado en su forma perfecta: मैंने देखा।

Según Gadre (1996) y Díaz (2023), los posesivos establecen una relación entre el objeto y la persona gramatical. Es de destacar que la selección del posesivo depende de la persona que posee el objeto, pero solo los posesivos antepuestos concuerdan en número (menos los pronombres *nosotros* y *vosotros* que también concuerdan en género) con el sustantivo y los posesivos pospuestos concuerdan tanto en número como en género.

Tabla 16: *Los posesivos antepuestos en español*

		<i>Singular</i>	<i>Plural</i>
<i>Un poseedor</i>	Yo	mi	mis
	Tú	tu	tus
	Él/ ella/ usted	su	sus
<i>Varios poseedores</i>	Nosotros/-as	nuestro/-a	nuestros/-as
	Vosotros/-as	vuestro/-a	vuestros/-as
	Ustedes/ Ellas/Ellos	su	Sus

Nota. Tomado desde Gadre (1997)

Tabla 17: *Los posesivos pospuestos en español*

	<i>Un poseedor</i>				<i>Varios poseedores</i>			
	<i>Singular</i>		<i>Plural</i>		<i>Singular</i>		<i>Plural</i>	
	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.
<i>Primera persona</i>	mío	mía	míos	mías	nuestro	nuestra	nuestros	nuestras
<i>Segunda persona</i>	tuyo	tuya	tuyos	tuyas	vuestro	vuestra	vuestros	vuestras

Segunda persona formal y Tercera persona	suyo	suya	suyos	suyas	suyo	suya	suyos	suyas
--	------	------	-------	-------	------	------	-------	-------

Nota. Tomado desde Gadre (1997)

Ejemplos para entender mejor las diferencias entre los posesivos antepuestos y pospuestos.

1. **Mi** hermano viene hoy.
2. Un hermano **mío** viene hoy.

La oración 1 representa el uso de un posesivo antepuesto. Esto significa que, según el contexto, solo tengo un hermano. Aunque, la posición de *un hermano mío* en la oración 2, implica el uso de un posesivo pospuesto y refiere a la posibilidad de tener más de un hermano.

Hindi

En hindi, se representan 2 ejemplos más para entender la diferencia:

3. *Mērā bhāī āj āēgā.* (मेरा भाई आज आएगा)

Mi hermano hoy vendrá.

4. *Mērā ēk bhāī āj āēgā.* (मेरा एक भाई आज आएगा)

Mi un hermano hoy vendrá.

La estructura de la oración 3 representa la presencia de un hermano solo (como en el caso de artículo definido en español).

En la oración 4, el posesivo interpolado no llega a ocupar la posición posnominal como en español. El posesivo en hindi ocupa la posición prenominal y detrás del artículo *ek*. El *ek* acá representa la cantidad de hermanos y puede que tenga más (dato implícito). Esta prueba rectifica las afirmaciones que se hicieron en el apartado de los artículos. En donde si *ek* fuera un artículo indefinido (categoría no determinante), no tendría la opción de figurar con el posesivo (categoría determinante) en el sintagma nominal. Siendo un numeral que fue la afirmación de Gadre (1996), puede figurar con el posesivo en el sintagma nominal y ambos constituyentes: el numeral y el posesivo, ocupando la posición prenominal. La secuencia *mērā ēk bhāī* (mi un hermano) es idéntica a *mērē dō bhāī* (mis dos hermanos), mere *tīn bhāī* (mis tres hermanos), etc.

- *Mērā ēk bhāī.* ((मेरा एक भाई)

Mi un hermano.

Comparación entre hindi y español

A través de este análisis contrastivo, el hindi hablante puede darse cuenta de la siguiente regla: cuando en hindi el posesivo tiene $\emptyset$ en la estructura profunda, su equivalente en el español es el posesivo que ocupa la posición prenominal.

Indefinidos

Español

Según Gadre (1996) y Díaz (2023), los indefinidos son las palabras que marcan indeterminación y sus límites no son totalmente fijos. En español, en cuanto a la forma, se puede dividir los indefinidos en el paradigma nominal por la flexión de género y número: alguno, ninguno, otro, uno, todo, mucho, poco; solo por la flexión de número: bastante, tal y cualquier y sin flexión: cada. En el caso de los indefinidos masculinos *alguno* y *ninguno*,

la forma cambia según la posición en relación con el sustantivo. Es decir, si el indefinido es antepuesto, se cambia de *alguno* a *algún* y de *ninguno* a *ningún*. En otro caso, cuando el indefinido tiene una posición pospuesta, se mantiene la forma de *alguno* y *ninguno*.

Hindi

En hindi los indefinidos en el sintagma nominal son: *thōda*, *bahut zyada*, *aisa*, *har*, *sara*, *sab*, *kōi bhi*, *bahut adhik*, *paryāpt*, *kafi*, *kōī*, *kōi nhi aur* y *kuch*. Para entender los indefinidos correspondientes, se expone la tabla hecha por Gadre(1996) y actualizada en este trabajo. La tabla mencionada tiene la categoría neutra para mostrar elementos en hindi y español. Ya se sabe que no existe un género neutro en hindi, por eso la categoría no refiere a indefinidos neutros, sino a su función neutra. Por ejemplo, *koi* en sí, se refiere a *alguno* en español, así tiene una función neutra. También *koi* concuerda con sustantivos masculinos y femeninos cuando están pospuestos (*kōi ādmi* (algún hombre), *kōi ladki* (alguna chica)). Lo mismo sucede en el caso de *kuch*.

Comparación entre hindi y español

Tabla 18: Comparación de los indefinidos en hindi y español

	Hindi		Español	
	Singular	Plural	Singular	Plural
Masculino	<i>thōdā</i> / थोडा	<i>thōdē</i> / थोड़े	poco	pocos
Femenino	<i>thōdī</i> / थोड़ी		poca	pocas
Neutro	<i>thōdā</i> / थोडा		poco	
Masculino			tanto	tantos

Femenino	<i>Bahut zyada/ बहुत ज्यादा</i>		tanta	tantas
Masculino	<i>Aisa/ ऐसा</i>	<i>aisē / ऐसे</i>	tal	tales
Femenino	<i>Aisi/ ऐसी</i>			
Neutro	<i>pratyēk /har/ प्रत्येक / हर</i>		cada	
Masculino	<i>sārā / सारा</i>	<i>sārē / सारे</i>	todo	todos
Femenino	<i>sārī / सारी</i>		toda	todas
Neutro	<i>Sārā/ सारा</i>		todo	
Masculino	<i>Sab/ सब</i>		todo	todos
Femenino			toda	todas
Neutro			todo	
Masculino	<i>Kōi/ कोई</i> <i>Kuch/ कुछ</i>		alguno	algunos
Femenino			alguna	algunas
Neutro			alguien	
Masculino			uno	unos
Femenino			una	unas
Neutro			uno	
Masculino	<i>kōī bhī / कोई भी</i>	<i>Kuch bhii/ कुछ भी</i>	Cualquier	cualesquiera
Masculino	<i>bahut /बहुत</i>		mucho	muchos
Femenino			mucha	muchas
Masculino	<i>Bahut adhik/ बहुत अधिक</i>		demasiado	demasiados
Femenino			demasiada	demasiadas
Masculino	<i>Paryapt/ पर्याप्त</i>		bastante	bastantes
Femenino	<i>kāfi / काफ़ी...</i>			

Masculino	<i>kōī nahīn / कोई नहीं</i>	ninguno	ningunos
Femenino		ninguna	ningunas
Masculino	<i>Aur/ और</i>	otro	otros
Femenino		otro	otras

Nota. Tomado y adaptado desde Gadre (1997)

Numerales Cardinales

Español

Los numerales cardinales sirven para expresar cantidades. En los dos idiomas, español e hindi, se refieren siempre a un sustantivo, pero pueden usarse solos cuando ya está claro de qué estamos hablando. Por ejemplo:

- *Tengo **tres** bocadillos para comer después de almorzar.*
- *Yo tengo **dos**.*

En español los números de cero a diez no tienen un patrón para ayudar a los estudiantes a memorizarlos. De once a quince, se puede decir que siguen un patrón y después de dieciséis a veintinueve, se escribe todo en una sola palabra. De treinta a noventa y nueve todos los números se escriben en dos palabras separadas, con y en la mitad.

Tabla 19: *Números cardinales en español e hindi*

Arábigos modernos	Devanagari (números)	Devanagari (números en palabras)	Devanagari (en letras romances)	Números en español
0	०	शून्य	shunya	cero

1	१	एक	ek	uno
2	२	दो	dō	dos
3	३	तीन	tīn	tres
4	४	चार	cār	cuatro
5	५	पाँच	pāmc	cinco
6	६	छः	chh	seis
7	७	सात	sāt	siete
8	८	आठ	āth	ocho
9	९	नौ	nau	nueve
10	१०	दस	das	diez

Nota. Ilustración propia

Hindi

En hindi, los números cardinales del 0-99 no siguen un patrón explícito y los estudiantes deberían memorizarlos todos. Sin embargo, a partir del cien, los números empiezan a tener patrones más predictibles.

Comparación entre hindi y español

Se encuentra que hay algunos números que se parecen en los dos idiomas por la raíz indoeuropea, por ejemplo: dos, nueve, siete y diez entre otros números. También algunos números en devanagari tienen similitudes en forma con los números en arábigos modernos, como 0, 2, 3, 10 entre otros.

Agregativos

Los agregativos son palabras adjetivas que se forman con la flexión masculina indirecta plural en –o de cualquier número cardinal a partir del 2, por ejemplo:

Tabla 20: Agregativos en hindi

Número	Número en hindi	Agregativo en hindi	Traducción al español
2	do	dono	Los dos
3	tīn	tīno	Los tres
4	cār	cāro	Los cuatro

Nota. Elaboración propia a partir de Díaz (2023) y Gadre (1996)

Se usan agregativos en hindi para enfatizar que un número de personas o cosas forma la totalidad de un conjunto. En el mismo orden de ideas, solo el agregativo (dono) es irregular pero el resto se forma agregando solo –o. Por ejemplo: ambos o los dos.

Ejemplos:

1. Ambos muchachos van a la escuela.

ḍōnō ladakē skula jatē hain.

दोनों लड़के स्कूल जाते हैं।

2. Mis tres búfalas se escaparon.

Mēri tinō bhēns bhāg gai.

मेरी तीनो भैंस भाग गई।

Si las personas o cosas ya habían sido mencionadas, el agregativo funciona como sustantivo en los dos idiomas: hindi y español. Por ejemplo:

Abajo está el amor, arriba está Dios. En medio de *los dos* está todo.

Nēēchē ishq hai, upar rab hai. In ḍōnō kē bich mē sab hai.

नीचे इश्क है, ऊपर रब है। एन दोनों के बीच में सब है।

Numerales ordinales

Español

En español, los números ordinales expresan orden en una serie:

- *El primer⁷ libro de Dan Brown.*
- *Estoy en el cuarto semestre de la maestría.*

Según la Real Academia Española, en la lengua hablada, los ordinales solo se usan hasta 10°. Después de 10° se usan los cardinales. Estos numerales concuerdan en género y número con el sustantivo, pueden ir acompañados con otros determinantes, cuando van acompañado con un sustantivo, generalmente van antes y cuando ya está mencionado el sustantivo, se puede escribir solos.

- Las *primeras* gafas que llevé eran bonitas.
- Mi *primera* prima se llama Andrea.
- Hoy es el *primer* día de vacaciones.
- Esta fue la *primera* casa de mi abuelo. Y esta fue la *segunda*.

Tabla 21: Números ordinales en hindi y español

Números	En devanagari con letras romances	En devanagari	En español
0°	shunya	शून्य	Cero
1°	pēhla	प्रथम/पहला	Primero(masculino)
2°	dusra	द्वितीय/दूसरा	Segundo
3°	tisra	तृतीय/तीसरा	tercero
4°	chautha	चौथा	cuarto
5°	panchva	पांचवाँ	quinto

⁷ Refiere a anexo 1 para ver la apócope masculina de número ordinal *primero*.

6°	chahtha	छठा	sexto
7°	sātvā	सातवाँ	séptimo
8°	athva	आठवाँ	octavo
9°	nauva	नौवाँ	noveno
10°	dasva	दसवाँ	décimo
11°	gyarhva	ग्यारहवाँ	Undécimo, décimo primero
12°	barhva	बारहवाँ	Duodécimo, décimo segundo
13°	tērhva	तेरहवाँ	Décimo tercero
14°	chaudhva	चौदहवाँ	Décimo cuarto
15°	pandrhva	पंद्रहवाँ	Décimo quinto

Nota: Ilustración propia

Hindi

Los ordinales en hindi, en general, tienen un patrón predecible después de séptimo y pueden formarse regularmente agregándole el sufijo –va al número cardinal.

Los ordinales también se declinan en hindi como en español. En español los ordinales tienen la flexión en el género principalmente y concuerdan con los sustantivos. En cambio, en hindi, para las cuatro primeras formas ordinales (que terminan en –a) se usan las terminaciones normales de un adjetivo, es decir –a para el masculino directo singular –e

para los demás casos del masculino y para el plural –i para todo el femenino (Díaz, 2023).

Por ejemplo: *dusra* (segundo)

Tabla 22: Declinaciones de número ordinal dusra (segundo/a)

Dusra (segundo)	Caso	Masculino	Femenino
singular	directo	दूसरा (dusra)	
	indirecto		दुसरी (dusri)
plural	directo	दुसरें (dusrē)	
	indirecto		

Nota. Tomado y adaptado de Díaz (2023)

Relativos

En español, según la RAE, hay tres formas relativas que funcionan como determinantes.

En la siguiente tabla, se exponen las tres posibilidades con ejemplos.

Tabla 23: Relativos determinantes en español

El determinante	número	género	Ejemplo
cuanto/cuan	singular	masculino	Traté de entregar a María cuanto dinero [= todo el dinero que] llevaba en el bolsillo ⁸
		femenino	Cuanta leche pueda tomar
	plural	masculino	Cuantos amigos quieras
		femenino	Cuantas tareas necesites
cuyo ⁹	singular	masculino	Yo soy en cuyo pecho dormiste.
		femenino	Había acudido a él [...], cuya juventud y talante le iban a permitir mangonearlo a su antojo...
	plural	masculino	Se asomó a una ventana a través de cuyos cristales se podía ver la televisión...
		femenino	Espero que esa fidelidad alcance no solo a mis descendientes, sino a todos aquellos en cuyas manos pudiera caer...
artículo+ cual	singular	masculino	El cual modo quisiera que fuera tan fácil
		femenino	La cual manera de hablar ninguno entiende
	plural	masculino	En los cuales sucesos, no me alargo

⁸ Tomado desde *rae.es*

⁹ Tomado desde *rae.es*

femenino	Las cuales causas son las ideas
----------	---------------------------------

Nota. Tomado y adaptado de la RAE

Según el PCIC, los determinantes expuestos en la tabla 23, se usan en niveles intermedio- avanzado por la complejidad sintáctica en la creación de las oraciones. Igualmente, la última categoría de (artículo+ cual) no se usa en la actualidad en la vida cotidiana ni profesional. No obstante, se observa que el PCIC enfoca en los pronombres relativos en los niveles básico-intermedio y no de los relativos determinantes.

Hindi

Tabla 24: *Declinaciones de relativo jitna en hindi*

jitna (cuanto)	Caso	Masculino	Femenino
singular	directo	जितना (jitna)	जितनी (jitni)
	indirecto		
plural	directo	जितने (jitnē)	
	indirecto		

Nota. Elaboración propia

Tabla 25: *Declinaciones de relativo jiska en hindi*

jiska (cuyo)	Caso	Masculino	Femenino
singular	directo	जिसका (jiska)	जिसकी (jiski)
	indirecto		
plural	directo	जिसके (jiskē)	
	indirecto		

Nota. Elaboración propia

En hindi los determinantes relativos son de dos tipos – *jitna* y *jiska*. Se puede referir a las tablas para observar sus declinaciones. ‘*Jitna*’ en hindi tiene la misma función que ‘cuanto’ en español y ‘*jiska*’ lo mismo que ‘cuyo’. Abajo se elabora la equivalencia entre los dos relativos.

Tabla 26: *Determinantes relativos en hindi y español*

Español	Hindi (letras latinas)	Hindi (Devanagari)
---------	------------------------	--------------------

Cuanta leche pueda tomar	<i>Jitna dōōdh pina hai</i>	जितना दूध पीना है
Cuantos amigos quieras	<i>Jitnē mitr chahiyē</i>	जितने मित्र चाहिए
Cuyo amigo	<i>Jiska mitr</i>	जिसका मित्र
Cuya amiga	<i>Jiski sakhi</i>	जिसकी सखि
Cuyos libros	<i>Jiski kitabēin</i>	जिसकी किताबें

Nota. Elaboración propia

Comparación entre hindi y español

El relativo en los dos idiomas tiene la concordancia de número y género, es más, en hindi también se tiene la flexión de caso. Como se vio en español, el uso de algunos relativos no es común en la actualidad y, por lo tanto, un estudiante de español no tiene que saber todos los relativos determinantes en el idioma. Además, en PCIC se exploran los relativos desde el paradigma pronominal. Ahora bien, la referencia de los relativos de hindi podría ayudar a un estudiante de español hindi hablante en entender que el relativo *cuanto* y sus formas refiere a la cantidad y el *cuyo* y sus formas se refiere al sustantivo directamente. Ya que, tanto en hindi como en español, el significado de estos dos relativos es similar.

5.1.1.6 Orden sintáctico

El orden sintáctico en los dos idiomas: hindi y español, son diferentes y por lo tanto se expondrá primero el orden sintáctico básico de los dos idiomas, es decir la estructura de la oración como base y luego se expondrán las modificaciones y componentes gramaticales extras que puedan afectar a este orden. La siguiente explicación se hará en una tabla ([anexo 2](#)) para entender las diferencias y similitudes de una forma sencilla. Así, después de hacer este análisis, se encuentran las siguientes posibles dificultades durante el aprendizaje de español:

El orden sintáctico general, aunque no siempre, en español es $S + V + O$, pero en hindi es $S + O + V$. Los estudiantes deberían ser conscientes de este cambio en el orden en niveles básicos.

El concepto de la concordancia en el paradigma nominal es una novedad para estudiantes hindi hablantes. Aunque este concepto existe en hindi, los estudiantes no son conscientes de esta regla gramatical. Además, en hindi la concordancia de género se refleja en el verbo. Este cambio podría ser una dificultad.

El ergativo juega un papel importante en hindi y así nacen las declinaciones. En el orden sintáctico se nota que a veces, el ergativo es equivalente a la preposición en español, pero no es el caso siempre. Esta diferencia es de destacar en la clase con hindi hablantes.

5.2 Análisis de errores

A finales de los sesenta y principios de los setenta, algunos investigadores intuyeron que el análisis contrastivo tenía limitaciones (Jain, 1974; Hammarberg 1974) porque hubo un gran número de errores que no se podían explicar desde la interferencia de la lengua materna del aprendiente. A partir de allí, las investigaciones en *análisis de errores* (AE) adoptaron la metodología de Corder que se había publicado en 1967. Esta metodología parte del análisis de la producción oral y/o escrita del estudiante en la lengua que está aprendiendo (la lengua materna no juega un papel importante). La teoría de análisis de errores de Corder se fortaleció más cuando Chomsky y Piaget (1979), proponentes del cognitivismo psicológico, empezaron a interesarse por el comportamiento humano y consideraron que el conocimiento es un proceso a través del cual el aprendiente/ individuo construye la realidad que le rodea. Ante esta novedad, es lógico preguntarnos la importancia de AE para las futuras investigaciones. Burt, Bulay y Krashen (1982, 1938, 1939) mencionan las ventajas de este modelo: supone una contribución a la lingüística aplicada (LA) (ya que, los errores son vistos como algo positivo y una indicación de aprendizaje), motiva a los profesores a ser conscientes de las dificultades y facilidades que tiene un estudiante, establece una jerarquía de esas dificultades, produce material de

enseñanza y actualiza métodos de evaluación para determinados objetivos y niveles. No obstante, como todas las teorías de análisis de errores, esta teoría de Burt, H. Dulay y Krashen tampoco carece de imperfecciones, ya que, en cuanto a la categorización de los errores, se observa una falta de precisión y especificidad motivada, sobre todo, por la confusión entre los errores interlingüísticos e intralingüísticos.

El modelo ha evolucionado bastante durante los años, de predecir y explicar tipologías de errores basados en taxonomías gramaticales, se pasó a estudiar los errores con una función más comunicativa. Esta actualización también se puede ver en el pensamiento de Corder quien publicó su trabajo acerca de AE en 1967. En este, el autor se enfoca en el inventario de los errores más frecuentes y en años posteriores (1971, 1981), amplió los objetivos hacia la práctica, enfocándose en la ampliación didáctica del mismo análisis y proporcionando a los alumnos suficientes herramientas para comunicarse satisfactoriamente. J.C Richards (1974) señala que AE puede utilizarse como fuente de corroboración de la información mediante el AC, el cual funciona como predictor de errores.

5.2.1 El concepto del error

En este apartado se reflexiona acerca del concepto del error formulado por S.P. Corder (1967). Él menciona la importancia de aclarar la diferencia entre una falta y un error porque fundamentalmente la falta no se considera en el AE, es porque la falta no es una producción consciente y sistemática que se repite y tiende a fosilizarse, en cambio, el error tiene esta tendencia.

Falta- Error de producción

Error- Error sistemático que afecta a la competencia del aprendiente y que es propiamente idiosincrásico.

Luego, Norrish en el año 1983 incluye el concepto de *lapsus* a la propuesta de Corder:

Falta- una desviación inconsciente y eventual.

Error- una desviación sistemática.

Lapsus- una desviación debido a factores extralingüísticos

A continuación, Corder (1967) y Maldonado (2013) definen la falta, el error y el lapsus de la siguiente manera.

- El lapsus se define como un incidente no planificado y, por lo tanto, no sistemático en el uso del lenguaje, que no indica un defecto en el conocimiento del propio idioma. Este fenómeno puede ser resultado de una falla en la memoria, estados físicos o condiciones psicológicas, como una emoción intensa.
- La falta en el desempeño lingüístico se distingue por revelar una competencia transitoria, pero a diferencia del error, se caracteriza por su falta de sistematización.
- Por otro lado, el error, al igual que la falta, revela una competencia transitoria, pero se diferencia de esta última por su naturaleza sistemática. Esta sistematicidad permite reconstruir la competencia transitoria, por lo que el hablante ve el error como un signo positivo que marca las diferentes etapas del aprendizaje y señala el proceso realizado.

5.2.1.1 El error considerado como síntoma de fracaso y síntoma del progreso

El error, considerado síntoma de fracaso, nació del estructuralismo (concepción conductista del aprendizaje), basado en la creación de nuevos hábitos y la exclusión de hábitos anteriores, o sea, la lengua materna se consideraba fuente de errores a eliminar (Saussure, 1971). Bajo este criterio, el estudiante era el responsable del tratamiento del error y el rol del docente era corregir y erradicar los “malos hábitos” de los estudiantes. Este tratamiento del error convenció a los estudiantes de que los errores eran culpa suya y resultó en el miedo al ridículo ante los demás. Así, generó un comportamiento como “aprendizaje defensivo” especialmente en las culturas asiáticas, en donde los conceptos de ridículo y perfección llegan a ser bloqueadores en las producciones escritas y orales.

En una visión contemporánea y positiva, basada en la idea de Chomsky (1965) de que el estudiante está creando su hipótesis durante la adquisición de la lengua, el error se considera natural, próximo a la realidad y como pruebas de esta hipótesis hasta recibir la corrección o la confirmación de su acierto. Estos errores son las muestras de fortalecimiento y reestructuración de la lengua. Bajo este criterio, el profesor juega un papel importante, por lo tanto, dirige a los estudiantes en el aprendizaje de la lengua con base en el AE y el entendimiento de las necesidades de los estudiantes.

Después de ver el error como síntoma de fracaso y de progreso, es importante tener en cuenta el método de corrección de errores. Es de destacar que muchas de las investigaciones que se presentaron en el estado del arte nos arrojan información sobre errores y sus frecuencias, pero el tratamiento de estos errores sigue siendo la decisión del docente. Desde la LA se considera que lo más importante es adecuar el error a la realidad de la clase, de la personalidad del alumno, de su tipo y la frecuencia con la que lo hace. Este método, en primer lugar, es un tratamiento reflexivo y toma en cuenta el contexto y, en segundo lugar, no le hace responsable al estudiante, sino le ayuda a reflexionar en el porqué del error.

Las dos metodologías (AC y AE) tienen el mismo objetivo: analizar la interlengua del estudiante. Sin embargo, hay dos posturas contrarias con respecto a la implementación de AC y AE. Hay un grupo de lingüistas que consideran al AC inadecuado y fragmentario y, por lo tanto, debería ser remplazado por el AE completamente (Buteau, 1970; Svartvik, 1972; Jain, 1974; Salma, 1971; Zydatiss, 1974; Hammarberg, 1974; entre otros); y hay otro grupo que ve el AE como un complemento de AC y considera que hay que trabajar paralelamente con ambos modelos para obtener resultados completos (Wilkins, 1969; Duskova, 1969; Nickel, 1972; Fisiak, 1981; Sajavaara, 1981; Marton, 1981; entre otros). El argumento del primer grupo, que el AC es fragmentario e inadecuado, no es motivo

suficiente para invalidar este modelo, por lo tanto, en nuestra investigación se toman en cuenta los dos modelos con AC como una primera etapa de análisis y luego AE para complementarla.

5.3 La interlengua

Como se vio al principio del apartado, el concepto de la interlengua (IL) surgió después de AC y AE en 1969 por Selinker, reelaborado posteriormente en 1972. Este nuevo campo nos lleva a considerar los modelos de la glosodidáctica del aprendizaje; AC, AE y IL, como etapas sucesivas de una aproximación al estudio global del lenguaje del estudiante en su L2.

L. Selinker (1972, 214) define el concepto de Interlengua como: “Un sistema lingüístico separado sobre cuya existencia podemos hacer hipótesis en el “output” de un estudiante al intentar producir la norma de la lengua meta.”

Desde la perspectiva de la LA, se han abordado tres perspectivas de la IL con diferentes denominaciones: interlengua (psicolingüístico) Selinker (1972), dialecto idiosincrásico (sociolingüística) Corder (1972a) y sistema aproximativo (lingüístico) Nemser (1971). Sin embargo, actualmente la decisión de varios lingüistas de usar el término Interlengua es unánime.

Ahora bien, esta IL del estudiante va incrementando en su complejidad bajo un proceso creativo durante la adquisición de la lengua meta. En la Figura 4, León (2019) hace una representación de la IL y la producción de errores a lo largo del desarrollo de la misma. De esta forma, León también desarrolla la idea de que la IL es un proceso dinámico por su carácter transitorio, que además se puede analizar en un estudio longitudinal.

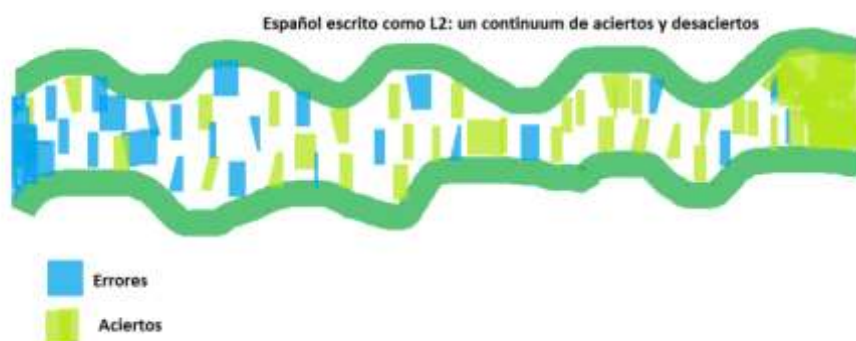


Figura 3: Los errores como parte de la Interlengua. *Tomado de León (2019)*

A continuación, se enumeran las características de la IL según algunos expertos (Alexopoulou, 2005; Baralo, 2004; Muñoz-Busols et. Al.,2016)

- a. Fossilización: Fenómeno lingüístico y la tendencia inconsciente de conservar las reglas de la L1 en la IL. Selinker (1972) durante su trabajo con el concepto de la IL estableció este término. La fossilización representa la estabilización de los errores lingüísticos, tal vez por la falta de corrección durante los niveles iniciales. Se considera que un error está fossilizado cuando el aprendiente no logra autocorregir el error a pesar de obtener correcciones y retroalimentación.
- b. Transferencia: Krashen (1981) introdujo este término en su teoría de la adquisición de segundas lenguas. Junto con Gass (2013) definieron la transferencia como la influencia de la lengua materna u otras lenguas aprendidas en la adquisición y aprendizaje de segundas lenguas o lenguas extranjeras. Se refiere a la interferencia unidireccional (de lengua materna a lengua meta) que puede generar un efecto positivo, facilitando el aprendizaje, o un efecto negativo, generando errores.
- c. Permeabilidad: Tendencia de permitir las reglas de la L1 la IL, que origina sobregeneralizaciones de las reglas. Ellis (1994) y Mackey (2004) abordaron el tema de la permeabilidad y lo definieron como la tendencia de los errores de una lengua a afectar la

producción en otra lengua. Se refiere a la interferencia bidireccional entre la lengua materna y la segunda lengua.

d. Variabilidad: Evidencia de que la IL es transitoria. Ortega (2014) ha trabajado el factor de la variabilidad en la producción lingüística considerando su impacto en la comprensión de la adquisición de segundas lenguas. Ella menciona que la variabilidad se refiere a las diferencias en la producción lingüística de un aprendiz en diversos contextos o momentos. Puede ser influenciada por factores como la motivación, la fatiga o la complejidad de la tarea.

5.4 Las taxonomías para el análisis de los errores

En este apartado se exploran algunas taxonomías ya presentes para analizar los corpus de aprendientes y las tendencias de investigaciones, para luego elegir y/o adaptar estas para el análisis de nuestro corpus CAELE/2¹⁰.

Del mismo modo, para realizar todo el proceso de determinar la frecuencia de los errores, identificar las tendencias de uso erróneo y evaluar el estado de la interlengua, se necesita un sistema de anotación. Por lo tanto, el CEA (computer aided error analysis, CEA- en inglés) se basa en diseñar y/o adaptar una taxonomía de las que ya existen para reconocer y etiquetar los errores en diversos niveles de categorización. Es conveniente explicar que el análisis de error asistido por computador es una metodología semiautomática que usamos para etiquetar los errores bajo ciertos criterios y categorías que dependen de las características de los errores y el alcance del estudio. Este etiquetado se usa para anotar los errores con el fin de identificarlos, clasificarlos, describirlos y explicarlos. Una vez terminado el proceso, se pueden observar las desviaciones en la

¹⁰ El subcorpus CAELE/2 Extranjeros Escrito está compuesto por muestras escritas de aprendientes de español como lengua extranjera. El subcorpus hace parte del proyecto *Corpus de Aprendientes de Español como Lengua Extranjera y Segunda (CAELE/2)* cuyo objetivo es "diseñar, recopilar, sistematizar y construir un corpus escrito y oral de muestras de aprendientes de español.

interlengua (según los criterios de Corder (1967) del aprendiente y obtener los resultados según las variables definidas en la investigación.

Con referencia a las taxonomías disponibles, vale señalar que Corder (1967), Alexopoulou (2006), De Alba Quiñones (2009) produjeron la base teórica de casi todas las taxonomías que se consideran en esta investigación. A su vez, Negrillo (2007) sustenta la importancia de las taxonomías y su utilización para evitar la subjetivación de los resultados de corpus bajo análisis. De acuerdo con la literatura encontrada, las primeras taxonomías se basaban en los siguientes tres criterios principales:

Tabla 27: *Ferreira & Elejalde (2020)*

Criterio descriptivo	Formas y estructuras presentes en la conformación de la interlengua. Las categorías son- la adición, omisión, falsa elección la elección errónea, la sustitución o forma errónea
Criterio lingüístico o gramatical	Delimita el nivel del error. Por ejemplo- a nivel de la palabra, la oración o el párrafo. Se relaciona con las características del error en sus formas gramaticales, ortográficas, sintácticas, semánticas, pragmáticas o discursivas de la lengua.
Criterio etiológico	Se basa en la transferencia negativa de la lengua materna del aprendiente. Interlingüístico o Intralingüístico.

Nota. Criterios principales para la creación de una taxonomía para el análisis de error.

Tomando en cuenta la tabla 19 y la existencia de estas 3 categorías implícita y explícitamente, se exploran las taxonomías de Santos, (1993), Alexopoulou (2005), Durao (2005), Ferreira (2016) y Ferreira y Elejalde (2020) para observar las similitudes y las modificaciones que se han dado en el análisis contrastivo.

Tabla 28: *Tipología de errores propuesto por Santos Gargallo (1993)*

Descriptivo	Pedagógico	Etiológico-lingüístico	Gramatical	Comunicativo
• Omisión • Adición • Formación errónea • Ausencia de orden oracional	• errores transitorios • errores sistemáticos	• errores interlingüísticos • errores intralingüísticos	• Fonológicos • Ortográficos • Morfológicos • Sintácticos • Léxicos • Semánticos • Pragmáticos	• locales • globales

Nota. Tomado de Gargallo (1993, 130)

Para ilustrar la primera propuesta (ver *tabla 28*), el autor propone una taxonomía elaborada con los criterios etiológico- lingüístico y gramatical, separados para profundizar más en los errores. Luego agregó criterios pedagógicos y comunicativos. Estas últimas categorías son una adición a la perspectiva de la *tabla 27* y pretenden analizar errores desde una perspectiva pedagógica, sociocultural y la interferencia de la lengua materna mediante el criterio etiológico.

Tabla 29: *Tipología de errores de Alexopoulou (2005)*

Criterio descriptivo	Criterio lingüístico	Criterio etiológico
Adición (AD) Omisión (OM) Elección falsa (EF) Forma errónea (FE) Colocación falsa (CF)	Nivel morfosintáctico: Artículos (ART) Determinantes Pronombres (PR) Verbos Adverbios (ADV) Preposiciones (PREP) Conjunciones (CONJ)	Interlinguales Intralinguales -Reducción a un sistema simplificado - Generalización

Nota. Tomado de Alexopoulou (2005)

La segunda taxonomía es la propuesta por Alexopoulou (2005), en la que menciona los criterios generales como el descriptivo, el lingüístico y el etiológico. El criterio lingüístico se ha delimitado a nivel morfosintáctico y el etiológico con más explicaciones intralingüales. De la misma forma, Duraó (2007), en la *tabla 29*, usa los mismos criterios como en la tipología propuesta por Gargallo (1993) pero con más clasificaciones dentro de cada criterio. Especialmente, en el criterio etiológico, en donde Gargallo (1993) usa solo dos clasificaciones: intralingüísticos e interlingüísticos, mientras que, Duraó (2005) menciona subclasificaciones de cada una. Por ejemplo, los errores interlingüísticos son de producción, subproducción, superproducción, mala interpretación, hipercorrección, mientras los errores intralingüísticos son de simplificación, hipergeneralización, inducción y producción excesiva. Los otros criterios también se dividen en otras clasificaciones que arrojan una explicación profunda sobre los resultados y las posibles razones de los errores cometidos.

Tabla 30: *Tipología de errores propuesta por Duraó (2005)*

criterio lingüístico	criterio gramatical	criterio etiológico	criterio pedagógico	criterio comunicativo
• adición • omisión • falsa colocación (ausencia de orden natural) • falsa selección (malformación de palabras o estructuras)	• fonético-fonológico • ortográfico • morfológico • sintáctico • léxico-semántico • discursivo	errores interlingüísticos: • de producción, • subproducción • superproducción • mala interpretación • hipercorrección errores intralingüísticos: • de simplificación • hipergeneralización • inducción • producción excesiva	• de comprensión • de producción • individuales • colectivos • orales • escritos	• global • local • irritante • por inhibición • por ambigüedad • sociocultural • estigmatizador • pragmático

Nota. Tomado de Duraó (2005, 143)

Hasta ahora la taxonomía más desarrollada es propuesta por Ferreira (2016) para la realización de los estudios de la interlengua en el corpus CAELE (*tabla 30*). El desarrollo

de esta taxonomía se hizo tomando en cuenta el análisis de muestras escritas de la interlengua de los extranjeros de diferentes países no hispanohablantes. Los dos criterios: lingüístico y etiológico respetan dos niveles de análisis: palabra y oración. Las categorías: gramática, léxico, ortografía funcionan a los tres niveles lingüísticos de las muestras escritas, asimismo es evidente la exclusión de la categoría fonética por su origen en la oralidad.

De acuerdo con la *tabla 30*, el criterio etiológico, se organiza considerando los aspectos provenientes de la lengua materna del aprendiente. Los errores interlinguales se distribuyen principalmente en transferencias directas o indirectas de la L1 u otras lenguas aprendidas hacia la LE, uso de traducción literal, analogías falsas, falsos cognados, calcos semánticos y cambio de código. Todas estas reflejan cómo el aprendiente usa una estrategia determinada para solucionar un problema de comunicación en la cual desconoce la norma o correspondencia de esta. (Ferreira, 2020).

Tabla 31: *Taxonomía para la clasificación de errores proyecto Fondecyt n.º 1110812 y n.º 1180974*

Criterio	Nivel	Categoría	Subcategoría		Adición	Falsa selección	Omisión
Lingüístico	Palabra/oración	Gramática	Categorías gramaticales		ADI-PREP	FI-PREP	OMI-PREP
			Coherencia textual	Concordancia sintáctica	FI-CG-GEN		
				Puntuación	ADI-COMB	FI-COMB	OMI-COMB
				Conectores	ADI-CONECT	FI-CONECT	OMI-CONECT
	Palabra	Léxico	Léxico creado por derivación		FI-LEX-CREADO-DERIV		
			Léxico morfología		FI-LEX-MORF		
			Léxico innecesario		ADI-LEX-ENNE		
		Ortografía	Acentual	ADI-ORTO-grave	FI-ORTO-grave	OMI-ORTO-grave	
			Literar	ADI-ORTO-LIT	FI-ORTO-LIT	OMI-ORTO-LIT	
			Dierética	ADI-ORTO-hiato	FI-ORTO-hiato	OMI-ORTO-hiato	
Etiológico	Palabra/Oración	Uso L1			HNO-L1		
		Uso L2			HNO-L2		
		Orden de palabras			orden de palabras		
		Analogía semántica			analogía semántica		
		Traducción literal			traducción literal		
		Sobregeneralización			sobregeneralización		
		Simplificación			simplificación		
Error de tipo					error tipo		

Nota. Tomado de Ferreira (2020, 5)

Para profundizar más en el criterio etiológico, Ferreira y Elejalde (2020) propusieron otra taxonomía como se ve en la *tabla 31*. Los errores interlingüísticos se refieren a la transferencia negativa y directa de la lengua materna u otras lenguas habladas por parte del aprendiente. Esta transferencia directa tiene 9 subclasificaciones: Las categorías gramaticales en el nivel de palabra, la concordancia sintáctica a nivel de oración, la estructura morfológica a nivel de palabra, el léxico a nivel de palabra, la coherencia textual a nivel de párrafo o texto, el cambio de código a nivel de palabra, los falsos cognados a nivel de palabra, la traducción literal a nivel de oración, la interferencia de otras lenguas aprendidas a nivel de palabra. En cuanto a la clasificación interlingüística el autor menciona seis tipos: la neutralización referida a la correspondencia dual entre uno o más usos de diferentes categorías gramaticales o significados de una palabra, la sobregeneralización de una regla de LE, la simplificación de una regla que se puede dividir en la aplicación incompleta, la aplicación incorrecta y el desconocimiento de la regla y por último, el léxico creado por derivación con la intervención de palabras de LE y LM.

Tabla 32: *Propuesta de taxonomía con criterio etiológico (Ferreira y Elejalde, 2020)*

Clasificación	Tipo	Profundidad	Nivel	Descripción LO
Interlingüístico	Transferencia directa	Categorías gramaticales	Palabra	Omisión, adición, falsa selección y forma errónea
		Concordancia sintáctica	Oración	
		Estructura morfológica	Palabra	
		Léxico	Palabra	
		Coherencia textual	Párrafo/texto	
		Cambio de código	Palabra	
		Falsos cognados	Palabra	
		Traducción literal	Oración	
		Interferencia de otras lenguas aprendidas	Palabra	

Intralingüístico	Neutralización			
	Sobregeneralización			
	Hipercorrección			
	Simplificación	Aplicación incompleta de la regla		
		Aplicación incorrecta de la regla		
		Desconocimiento de la regla		
	Léxico creado por derivación			

Nota. Tomado de Ferreira y Elejalde (2020, 8)

Las mismas autoras profundizaron en los errores interlingüísticos de transferencia directa de la LM del aprendiente con un énfasis en las categorías gramaticales en el nivel de la palabra u oración. Como se observa en la tabla 32, los errores gramaticales tienen las siguientes categorías: preposiciones, sustantivo, adjetivo, pronombre, verbo y artículos. La taxonomía abarca varios paradigmas como el verbal, nominal y preposicional con sus posibles etiquetas de omisión, adición, falsa selección.

Tabla 33: *Errores de transferencia de estructuras gramaticales*

Profundidad	Categoría	Etiqueta ejemplo
Uso incorrecto de las categorías gramaticales a nivel de palabra u oración	Preposición	OME-PREP-A
		FS-PREP-CON
		ADP-PREP-DE
	Sustantivo	OME-SUST
		FS-SUST
		ADP-SUST
	Adjetivo	OME-ADJ
		FS-ADJ
		ADP-ADJ
	Pronombre	FS-PRON-IMPER
		FS-PRON-relativo
		FS-modo-SUBJ/INDIC/IMPER
	Verbo	FS-tiempo verbal
		FS-formas-personales
		FS-verbo-AUX
	Artículos	OME-ART-DEF
		FS-ART-INDEF
		ADP-ART-DEF

Nota. Tomado de Ferreira y Elejalde (2020, 6)

Los criterios describen el error desde un punto de vista que puede ser una visión pedagógica, comunicativa, etiológica, lingüística entre otras. También hay estudios interdisciplinarios con varios enfoques que brindan más información sobre el error cometido y su posible didactización en el futuro. Estas taxonomías servirían de base para cumplir con los objetivos de la investigación y el análisis de corpus CAELE/2 con la población meta.

Con el propósito de cumplir con los objetivos de esta investigación, se realizará un primer etiquetado para elegir el paradigma con máximo número de errores y adaptar una taxonomía. Los factores destacados para decidir la taxonomía para el análisis de errores es la frecuencia de los errores y el objetivo de la investigación y por eso conviene poner énfasis en el desarrollo, la elección y la adaptación de la misma.

5.5 La lingüística de corpus

Ha sido debatido si la LC es una teoría o una metodología por la involucración de tecnologías informáticas en el asunto (Svartvik, 1992; McEnery & Wilson, 1996; Kennedy, 1998; Stubbs, 1996, 2001; Bonelli, 2001), que ha sido una fortaleza para muchos, pero una debilidad para otros. En la actualidad, la LC tiene varias implicaciones metodológicas para el estudio de la lengua, como: análisis, descripción y enseñanza de lenguas. Así, LC tiene aplicaciones múltiples que ofrecen una aproximación específica de los datos lingüísticos de la lengua (Parodi, 2008).

De acuerdo con McEnery y Hardie (2012: 1-2) la lingüística de corpus (LC) se define como:

Set of machine-readable texts which is deemed an appropriate basis on which to study a specific set of research questions. The set of texts [...] is usually of a size which defies analysis by hand and eye alone within any reasonable timeframe.

La lingüística de corpus como campo de estudio ha evolucionado mucho en los últimos años, conviene destacar que el corpus en el siglo XXI es una recopilación de datos digitales para llevar a cabo un análisis apropiado para una investigación específica. Inicialmente, este campo se desarrolló en la lengua inglesa por su mayor difusión internacionalmente pero hoy en día existen varios tipos de corpus en diferentes lenguas. Uno de estos es el corpus de aprendientes, que es una base de datos de las muestras reales y auténticas

(escritas u orales) de los aprendientes de una lengua. Por lo general, estas muestras reales son de la interlengua de los estudiantes, de la autenticidad y la dinamicidad de ella para llevar a cabo un análisis bajo los criterios de una investigación.

Así mismo, la lingüística de corpus también ofrece estudios transversales y longitudinales. Los estudios transversales, esencialmente, se componen de una recolección de datos, en este caso muestras orales o escritas, con el uso de la segunda lengua solo una vez. Estas recopilaciones muchas veces son preguntas abiertas con aprendientes de características similares, por ejemplo: la lengua materna, el nivel de la segunda lengua, el nivel educativo entre otras. Por otro lado, para un análisis longitudinal, se necesita un corpus que recopile datos de los aprendientes durante varios intervalos de tiempo durante el desarrollo de la L2. Esta metodología ofrece un análisis profundo sobre el desarrollo de la interlengua en diferentes niveles y etapas del aprendizaje de L2 con una población meta.

Ahora bien, el corpus de aprendientes arroja colecciones sistemáticas, auténticas y contextualizadas del uso de la segunda lengua o lengua extranjera. Esta recopilación se puede explotar con varias metodologías que brinda el mismo campo de la LC. Estas metodologías pueden ser de análisis cualitativo o cuantitativo, transversal o longitudinal. En cuanto al análisis cualitativo, se trata de observar al locutor, la descripción y la explicación de la segunda lengua o lengua extranjera en un ámbito social o cultural que revela los factores para el uso de la lengua. En el análisis cuantitativo, el método es, preferiblemente, deductivo, orientado al producto de análisis y diseñado para probar alguna hipótesis. Las muestras, por lo general, son identificables, clasificables, cuantificables y se pueden analizar sistemáticamente.

Autores como Parodi (2008), Hincapié y Bernal (2017) hablan sobre las ventajas de LC en la enseñanza de ELE. Existen varias posibilidades en la explotación de un corpus de

aprendientes con ciertas características en común, como la lengua materna, la segunda lengua, el nivel educativo, el motivo de aprender el español entre otros factores. Con base en estas características comunes se pueden investigar problemas específicos que tienen los estudiantes de ELE. También se puede explotar este corpus como un material en el aula. Esta explotación, por lo general, es la didactización de este corpus dependiendo de la frecuencia de los errores y el tema abordado en el currículo. Así, los profesores pueden darse cuenta de los posibles errores y las fosilizaciones comunes entre los estudiantes que tienen la misma lengua materna, nivel educativo, nivel de lengua meta entre otros factores.

6.Marco metodológico

6.1 Diseño, alcance y paradigma

El presente trabajo tiene su base en la metodología mixta. Los métodos mixtos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos y cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada (metainferencias) y lograr un mayor entendimiento del fenómeno bajo estudio (Hernández-Sampieri y Mendoza, 2008). Chen (2006) los define como la integración sistemática de los métodos cuantitativo y cualitativo en un solo estudio con el fin de obtener una “fotografía” más completa del fenómeno y señala que estos pueden ser conjuntados de tal manera que las aproximaciones cuantitativa y cualitativa conserven sus estructuras y procedimientos originales (“forma pura de los métodos mixtos”); o bien, que dichos métodos pueden ser adaptados, alterados o sintetizados para efectuar la investigación y lidiar con los costos del estudio (“forma modificada de los métodos mixtos”).

En ese orden de ideas, esta metodología mixta es de multimétodos que brinda un panorama completo de los resultados. De la misma forma, se usa el diseño explicativo

secuencial (DEXPLIS), lo que quiere decir que primero se analizan los datos cuantitativos y luego los cualitativos. La *figura 5* explica el diseño adaptado de Hernández; Fernández; Baptista (2014), la primera fase como recolección de datos cuantitativos que en el presente estudio se dio bajo el proceso de la recopilación de corpus y el perfil lingüístico de los participantes, la segunda fase es la recopilación de datos cualitativos de las muestras escritas por parte de los participantes. Después, para la interpretación del análisis, primero se hizo la sistematización de datos cuantitativos y cualitativos y al final se hizo una discusión exhaustiva en el apartado de los resultados.

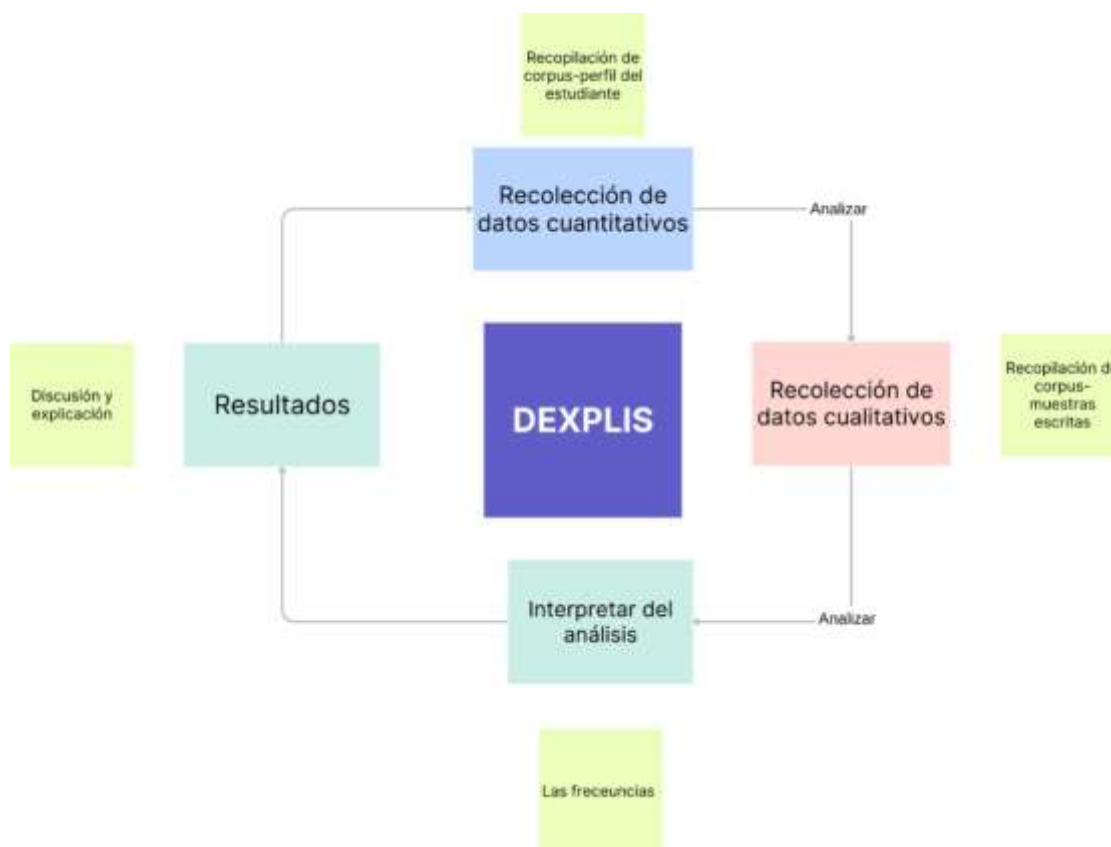


Figura 4: Esquema del diseño explicativo secuencial adaptado de (Hernández y Mendoza, 2014)

Asimismo, esta investigación tiene un alcance descriptivo ya que busca describir fenómenos, situaciones, contextos y sucesos; esto es, detallar las frecuencias y cómo se manifiestan los errores. Por tanto, se busca especificar las propiedades, las características y los perfiles de los participantes para darle contexto al análisis de error y por eso, se explota

la IL como una metodología, en donde, se describe el perfil del estudiante (un estudio transversal). En función de lo planteado, las muestras escritas recopiladas de hindi hablantes en el corpus CAELE/2, usuarios de español como lengua extranjera, se considerarán como la interlengua de los estudiantes.

Por último, esta investigación se basa en la metodología de la lingüística de corpus y usa el corpus de estudiantes indios recopilado y disponible en CAELE/2. El análisis de este corpus se basó en los pasos mencionados por Gargallo (1998). Es importante destacar que, en esta investigación, por el mismo alcance del trabajo, la falta de la literatura en los estudios del corpus en la India y la limitación del tiempo, se abordó solo el análisis a nivel morfosintáctico en el paradigma nominal de la lengua española y el hindi. De tal forma, abajo elaboramos los pasos adaptados y modificados (desde Gargallo (1998) para este proceso:

- a. Descripción morfosintáctica de español e hindi en el paradigma nominal;
- b. Cotejo de las descripciones de español e hindi;
- c. Elaboración de un listado preliminar de estructuras no equivalentes;
- d. Predicción y descripción de las dificultades en los resultados;
- e. Jerarquizar las dificultades según los resultados.

6.2 Muestra y población

En esta investigación se recopilaron las muestras de 40 participantes indios en total, como se ha mencionado antes, se eligieron a 23 estudiantes hindi hablantes. En total se analizaron 59 muestras, de las que 19 eran de nivel A2, 13 de nivel B1 y 27 de nivel B2. El grupo de participantes lo componían 17 estudiantes mujeres y 6 estudiantes hombres, con edades entre los 19 y los 57 años. Todos eran hablantes de hindi como lengua materna con otras segundas lenguas y lenguas extranjeras. Las segundas lenguas comunes eran inglés, punyabí, tamil, haryanvi, bengalí, marathi, urdu, bodo, kannada y telugú. Las lenguas

extranjerías eran español, chino mandarín y coreano. Cabe señalar que todos hablaban inglés como lengua materna o segunda lengua, dado que es una lengua que hace parte de las políticas lingüísticas del sistema educativo bilingüe de la India. El español era otra lengua común que todos hablaban y ya llevaban entre un año y medio a 4 años de aprendizaje. 11 participantes tenían experiencia de inmersión cultural en países hispanohablantes como en Colombia y España. 20 de ellos tenían contacto con el español por las redes sociales, amigos hispanohablantes, viajes que hicieron, las películas y la música en español. También, mencionaron que el modo subjuntivo era el tema que se les dificultaba hasta la actualidad y otros mencionaron temas como las diferencias entre *por* y *para*, la construcción de la oración y los pretéritos. Mencionaron específicamente la ayuda externa que tomaron durante el desarrollo del formulario: el diccionario *wordreference*, *spanishdict* y otros diccionarios disponibles en internet.

6.3 Instrumentos

Se utilizaron dos instrumentos principales: el cuestionario y la taxonomía adaptada. El cuestionario para la recopilación de corpus tiene los aspectos cualitativos y cuantitativos y se usó para recopilar todos los metadatos de los participantes y las muestras escritas. En cuanto a la taxonomía usada en la investigación, se usó primero la taxonomía creada por Pico (2022) para el análisis preliminar y luego se adaptó la taxonomía de Ferreira y Elejalde (2020) para el análisis en el paradigma nominal en las muestras escritas.

6.3.1 Cuestionario (Recopilación de datos en la población india)

El instrumento principal para la recolección de datos es un formulario de *Google* que se adaptó del instrumento de recolección de muestras escritas y orales ya desarrollado y validado por el Instituto Caro y Cuervo en el subcorpus CAELE/2 (Hincapié et al., 2018)). Es de resaltar que ya se había validado el instrumento y por eso, para esta investigación, no se repitió el proceso porque solo se actualizaron las preguntas del cuestionario. Ahora bien,

este formulario está dividido en 3 partes y recoge primero el perfil lingüístico, luego las muestras escritas y termina con una firma del consentimiento informado. Para mantener su formato autónomo y una mejor comprensión, las instrucciones están escritas en inglés y español, especialmente para los aprendientes de niveles básicos.

La primera parte se compone del título y la descripción de la investigación. Es un formulario que se diligencian de forma autónoma y por eso, se incluyen algunas recomendaciones, por ejemplo: llenar el formulario usando un portátil para facilitar la lectura y la escritura. En el mismo apartado se pide el perfil lingüístico a los estudiantes, entre los cuales tienen que mencionar la edad, el sexo, la nacionalidad, el país de nacimiento, el país donde vive ahora, la/s lengua/s materna/s, la/s segunda/s lengua/s, la/s lengua/s extranjera/s, el nivel educativo, la profesión, el nivel de la lengua española, el tiempo que lleva estudiando el español, alguna experiencia de inmersión cultural en países de habla hispana, la institución donde aprendió el español y el contacto con este idioma. También se pregunta por el tema más difícil y el tema más fácil en español para ellos. Todos los detalles sobre el perfil lingüístico del aprendiente tienen la función de delimitar el AE para que se enfoque en una población con características lingüísticas similares. En esta investigación, se delimitará el análisis al nivel de español y la lengua materna hindi. Del mismo modo, el resto de la información recogida respaldará al análisis cualitativo después de un análisis cuantitativo inicial.

En la parte 2 del formulario, primero se mencionaron las instrucciones para responder a las siete preguntas con algunas recomendaciones como la cantidad mínima de preguntas que deberían responder. Es de destacar que, por el diseño del formulario y la inversión del tiempo que se necesita para llenarlo, no se obligó a los aprendientes a responder a cierta pregunta o número de preguntas. Además, para facilitar y animar a los estudiantes, se puso la opción de teclear en el formulario mismo o escribir en el papel y subir la foto en la

plataforma. Las preguntas eran de tipo descriptivo, narrativo y argumentativo y adaptadas a contextos globales como las redes sociales, la educación, correos electrónicos y la inteligencia artificial, entre otros. Todas las preguntas mencionadas en el instrumento diseñado por los investigadores del Instituto Caro y Cuervo sirvieron para este cuestionario, pero al ser propuestas hace 5 años, no estaban contextualizadas a la realidad actual. Por lo tanto, se actualizaron al contexto actual, especialmente las preguntas relacionadas con las redes sociales y temas políticos. En cambio, se pusieron las preguntas con contexto global como la inteligencia artificial y el impacto positivo y negativo de las redes sociales. En la *tabla 34* se menciona la pregunta, la categoría y el tipo de pregunta.

Tabla 34: Preguntas para recoger las muestras escritas

Pregunta	Categoría	Tipo
Escríbele un correo a un amigo contándole cómo es un día en donde vives o en una ciudad a la que hayas viajado.	Correo	Descriptivo y narrativo
Quieres hacer amigos en una red social. Escribe un perfil que incluya información personal.	Redes sociales	Descriptivo
Describe como fue tu experiencia en un hotel o restaurante y recomienda o no este lugar.	Experiencia en un hotel o restaurante	Descriptivo
Redacta un texto en el que expliques cómo era la educación antes, cómo es ahora y cómo crees que será en los próximos años.	La educación	Narrativo
Redacta una queja o reclamo a una empresa de telefonía celular por su mal servicio, describe la situación, justifica el reclamo e intenta proponer alguna solución.	Reclamo/queja	Narrativo y argumentativo
¿Qué interpretación le das a la imagen?, ¿Con qué problemática social la relacionas?, ¿Qué soluciones darías a esta problemática?	Redes sociales	Argumentativo
Lee el siguiente fragmento. ¿Qué opinas de la propuesta de la inteligencia artificial y ChatGPT? Argumenta tu posición.	Inteligencia artificial	Argumentativo

Nota: Elaboración propia

En el último apartado se incluyó el consentimiento informado con más detalles sobre la investigación, la ausencia de riesgos y una nota sobre la confidencialidad de los

participantes. El consentimiento pedía información de un documento legal en su país o en el extranjero y por último el formulario preguntándoles a los encuestados si usaron algún recurso o ayuda digital para responder a las preguntas, así como la retroalimentación sobre la experiencia con el formulario (más información en el [anexo 3](#)).

6.3.2 Taxonomía para la clasificación de errores

Las taxonomías usadas para el análisis de las muestras escritas fueron dos. La primera se usó para hacer un análisis preliminar y encontrar las frecuencias generales en el paradigma verbal, pronominal, nominal, preposicional y otros. Para realizar este proceso se usó la taxonomía propuesta por Pico (2022) que toma en cuenta el análisis de errores inicial de los sistemas gramaticales a nivel morfosintáctico. En la *tabla 35* se mencionan estas categorías y las subcategorías para observar las frecuencias iniciales de los errores y luego seleccionar la categoría con más errores.

Tabla 35: *Taxonomía general para el primer etiquetado de errores*

Sistema verbal	Concordancia sujeto verbo
	Conjugación (tiempo y modo)
	Objeto directo e indirecto
	Adverbios
Sistema pronominal	Valores del se
	Leísmo
	Pronombres relativos
	Selosismo
	Pronombres de objeto directo y objeto indirecto
Sistema nominal	Artículos definidos e indefinidos
	Concordancia de género y número de sustantivo y modificadores
	Determinantes: artículos, demostrativos, posesivos, indefinidos, numerales, cuantificadores.
Sistema preposicional	
Otros	Sistema conjuncional, adverbios que no modifican al verbo, etc.

Nota. Tomado de Pico (2022, 64)

La segunda taxonomía se desarrolló después de analizar las frecuencias de este análisis inicial y elegir el paradigma con máxima frecuencia de errores para luego, abordar este sistema con más profundidad. De este modo, se profundizó (descrito en los resultados) el

paradigma nominal por su alta frecuencia en el análisis preliminar. Para etiquetar los errores en el paradigma nominal, se adaptó la taxonomía de Ferreira y Elejalde (2020). Esta taxonomía tiene los siguientes criterios: lingüístico y etiológico. El criterio lingüístico explora los errores al nivel morfosintáctico en el sistema nominal y describe si el error es un error de adición, omisión, falsa selección o sustitución. Por otro lado, el criterio etiológico indaga sobre el porqué de los errores e intenta razonarlos desde las perspectivas interlingüística e intralingüística (tabla 36).

Tabla 36: *Taxonomía adaptada de Ferreira y Elejalde (2020)*

Criterio	Nivel	Categoría	Subcategorías		Descripción				
					Adición	Omisión	Falsa selección	Sustitución	
L I N G Ü Í S T I C O	P	M	Categorías gramaticales	Artículo definido	AD ART DEF	OM ART DEF	FS ART DEF	SUS ART DEF	
	A	O		Artículo indefinido	AD ART IND	OM ART IND	FS ART IND	SUS ART IND	
	L	R		Artículo neutro	AD ART NEU	OM ART NEU	FS ART NEU	SUS ART NEU	
	A	F		Determinante demostrativo	AD DET DEM	OM DET DEM	FS DET DEM	SUS DET DEM	
	B	O		Determinante posesivo	AD DET POS	OM DET POS	FS DET POS	SUS DET POS	
	R	S		Determinante relativo	AD DET REL	OM DET REL	FS DET REL	SUS DET REL	
	A	I		Determinante numeral	AD DET NUM	OM DET NUM	FS DET NUM	SUS DET NUM	
	E	T		Determinante indefinido	AD DET IND	OM DET IND	FS DET IND	SUS DET IND	
	N	A		Adjetivo	AD ADJ	OM ADJ	FS ADJ	SUS ADJ	
	O	X		Sustantivo	AD SUST	OM SUST	FS SUST	SUS SUST	
	O	I	Aspectos sintácticos	Orden sintáctico	ORD				
	R	S		Concordancia sintáctica	CONC GEN				
	Palabra		Morfología			AD MORF	OM MORF	FS MORF	SUS MORF
	Etiológico	Palabra en oración	Morfosintaxis	Interlingüístico	Transferencia de LM (hindi)		ETIO LM		
Simplificación					ETIO SIM				
Intralingüístico				Sobregeneralización		ETIO SOBRE			
				Hipergeneralización		ETIO HIPER			

Nota. Elaboración propia

Con el objetivo de constatar la fiabilidad del instrumento para realizar la clasificación y el análisis de errores, se desarrolló un instrumento de validación para revisar la validez de la taxonomía propuesta. Este documento de validación estaba compuesto por la taxonomía adaptada con los dos criterios: lingüístico y etiológico, la descripción de la taxonomía con ejemplos, los fragmentos para etiquetar con un ejemplo etiquetado, un espacio para comentarios extras y al final una rejilla de evaluación ([anexo 4](#)). La validez del instrumento se refiere al grado en que un instrumento mide la variable en cuestión y si mide lo que realmente debería medir, de acuerdo con *voces calificadas* (Hernández-Sampieri y Mendoza, 2008). Bajo esta referencia, se incluyó un apartado para etiquetar algunos fragmentos de las muestras escritas de CAELE/2 con la taxonomía adaptada de Ferreira y Elejalde (2020). Ahora, la fiabilidad del instrumento se refiere al grado de precisión que ofrece la taxonomía mediante un instrumento de evaluación. Esta necesidad de evaluar la fiabilidad se cumplió con la rejilla de evaluación al final del documento en donde el calificador externo tenía que evaluar ciertos criterios de escala 1 (el punto mínimo) a 5 (el punto máximo) para calificar el instrumento según el objetivo de la investigación. Los resultados de este proceso se presentan adelante en la descripción de las fases.

6.4 Herramienta para el análisis de datos (UAM corpus tool)

A partir de la búsqueda de las plataformas de análisis de errores, se encontraron diversas opciones que se emplean en la LC y para la anotación de las muestras escritas. En concreto: Max QDA, SPSS, entre otros, pero tienen ciertas limitaciones en el análisis de corpus porque no están diseñados específicamente para este tipo de datos, sino datos globales. Para solucionar esta desventaja, se eligió la plataforma *UAM corpus tool* por su diseño y los datos aprovechables que arroja durante un análisis cuantitativo. De tal forma, se encontró la investigación de Cabrera y Gómez (2017) para justificar el uso de este software. Vale destacar que, por su utilidad y la practicidad en el análisis de errores en un

corpus, el programa *UAM corpus tool* sirvió en este trabajo para anotar las muestras escritas de los aprendientes hindi hablantes en el corpus CAELE/2.

UAM corpus tool fue útil para anotar el corpus en las dos etapas del análisis, primero en el análisis preliminar y en la elección del paradigma con máxima frecuencia de errores ya que este programa está diseñado específicamente para analizar y anotar los corpus. Y, en segundo lugar, para anotar los errores nuevamente después de adaptar la taxonomía para el paradigma nominal. Una de las ventajas del programa es una anotación del mismo error en diferentes paradigmas, esto quiere decir que se podrá analizar un error a diferentes niveles como: primero el sistema verbal o el sistema nominal entre otros y luego, las categorías dentro de un sistema en específico.

6.5 Fases de la investigación

6.5.1 Recopilación de corpus

El primer paso para realizar este trabajo fue la recopilación de corpus de estudiantes indios, dado que, no había un corpus de aprendientes de español de India, se hizo una recopilación libre, es decir, el formulario se envió a todos los participantes indios con diversas lenguas maternas. Después de la adaptación del instrumento (ver [anexo 3](#)), se envió a más de 10 docentes y aproximadamente 12 instituciones por correo, además, a los contactos personales y estudiantes por las redes sociales. El proceso empezó con un pilotaje para comprobar el entendimiento del formulario (ya validado por el Instituto Caro y Cuervo) para su uso autónomo por parte del público general y, tras recibir retroalimentación positiva y el formulario diligenciado correctamente por los participantes, se recogieron los datos por aproximadamente 5 meses. Se considera importante señalar que hubo participantes que llenaron solo el perfil con sus metadatos y no las muestras escritas, por lo tanto, se decidió eliminar sus datos por la falta de aporte investigativo.

6.5.2 Sistematización de datos

La sistematización de datos comenzó tras la recolección de corpus de 28 estudiantes, aunque el formulario siguió abierto para más participación. Este proceso se realizó a través de *Microsoft Excel* y *Microsoft Word*. Con el objetivo de proteger la identidad de los participantes, se sistematizaron los perfiles en códigos según la *tabla 37*. De la misma manera, se codificó cada muestra escrita con la sistematización mostrada en la *tabla 38*.

Tabla 37: *Codificación para el perfil del estudiante*

Subcódigo	Correspondencia
ELE	Enseñanza de español como lengua extranjera
ME	Muestra escrita
0200	Número de participante en el corpus CAELE/2
M	Masculino o Femenino (Género)
B2	Nivel (A1-C2)
IND	Nacionalidad (India)
Código final	ELEME0200MB2IND

Nota: Elaboración propia

Tabla 38: *Codificación de la muestra escrita*

Subcódigo	Correspondencia
ELE	Enseñanza de español como lengua extranjera
ME	Muestra escrita
0220	Número de muestra en el corpus CAELE/2
Código final	ELEME0220

Nota: Elaboración propia

Conviene señalar que muchos participantes respondieron a más de una pregunta y por eso el número de estudiantes (39) no corresponde al número de muestras (81). De tal forma, para aclarar esta sistematización tanto para esta investigación como para las futuras investigaciones, se crearon 3 archivos ya disponibles en corpus CAELE/2 (*Tabla 39*).

Tabla 39: *Tipos de archivos en el corpus CAELE/2*

Tipo de archivo	Contenido
-----------------	-----------

TXT	La muestra escrita
WORD	El perfil lingüístico del participante y la muestra escrita.
PDF	El perfil lingüístico del participante y la muestra escrita.

Nota: Elaboración propia

6.5.3 Validación del instrumento

Se analizó la validez y fiabilidad de la taxonomía adaptada por el juicio de expertas, para luego poder generalizar los resultados de la investigación y aportar a futuros estudios. Para realizar el proceso, primero se adaptó la taxonomía, se revisó y se etiquetaron errores de algunos fragmentos de las muestras escritas (primer etiquetado). Segundo, se envió a dos calificadoras externas, un documento con la taxonomía adaptada, algunos fragmentos para etiquetar los errores y una rejilla de evaluación del instrumento. Tercero, se analizaron los dos documentos de la validación y se encontró que las dos calificadoras etiquetaron errores adecuadamente en el paradigma nominal, pero una de ellas también entró en el nivel ortográfico y así calificó la rejilla de la evaluación.

La primera calificadora dio 25 de 25 puntos en la evaluación del instrumento, la otra calificó 22 de 25. También mencionó que hubo errores que no pudo etiquetar desde ninguna subcategoría. Para solucionar el problema, se agregó la subcategoría de *morfología* y su respectiva descripción de adición, omisión, falsa selección y sustitución a la taxonomía. Después de estas modificaciones, se envió el instrumento a una tercera calificadora. Ella calificó el instrumento 24 de 25 porque no pudo etiquetar un fragmento. Aquí cabe aclarar que hubo un fragmento en el instrumento que no tenía ningún error en el sistema nominal. Por tanto, se considera que el instrumento cumplió con su función y logró etiquetar todos los errores.

Al final, para revisar la fiabilidad del instrumento, se comparó el nivel de las coincidencias entre los errores etiquetados entre las tres calificadoras. Se encontró que,

entre las calificadoras había una coincidencia de aproximadamente 78% y a la luz de las observaciones y comentarios de ellas, se agregó la categoría de la morfología a la taxonomía, dado que la taxonomía no estaba cumpliendo con su función de etiquetar todos los errores posibles en el paradigma nominal. Así, la tercera calificadora etiquetó casi todos los errores e hizo algunos comentarios acerca de la corrección en cuanto a la forma que se refleja en el ([anexo 4](#)). Finalmente, con los resultados de los tres documentos, se pudo validar el instrumento para anotar todas las muestras escritas bajo consideración en esta investigación.

6.5.4 Análisis de metadatos

Después de la sistematización de datos de todos los estudiantes, se analizaron sus metadatos para elegir a los participantes y muestras pertinentes a esta investigación. Esta decisión se tomó por los criterios ya establecidos a partir del diseño, el alcance y el paradigma de esta investigación. Por lo tanto, se eligió a los 23 estudiantes de habla hindi para hacer un análisis contrastivo y análisis de errores de la interlengua. La información sobre la población meta está mencionada en el apartado: Muestra y población.

6.5.5 Criterios para el análisis de datos

El análisis inicial empezó con la taxonomía propuesta por Pico (2022) con las categorías de la *tabla 35*. Después del análisis preliminar, se evidenció la cantidad de muestras escritas en cada nivel y se decidió descartar las muestras de niveles A1 y C2 porque el número de las muestras en estos niveles no era representativo (A1 con 2 muestras de 3 líneas y C1 con 6 muestras de un párrafo). Dadas estas razones, se decidió analizar los tres niveles: A2, B1 y B2 y no solo B2 (con el mayor número de muestras escritas). En segundo lugar, los tres niveles tienen números suficientes para considerarlos

como muestra representativa de cada nivel¹¹. En tercer lugar, los errores encontrados en los tres niveles dieron evidencias sobre el avance y las fosilizaciones durante el aprendizaje de español, lo que se discutirá en el apartado de los resultados y la discusión.

6.6 Consideraciones éticas

En esta investigación empírica, el trabajo de campo se hizo de forma virtual y se logró recopilar el corpus autónomamente por parte de los estudiantes. Así, para mantener y seguir las normas éticas de la comunidad investigativa (Helsinki, 1964, 2008), se creó y se incluyó un consentimiento informado al final del formulario en el que los participantes tenían que anotar el número de documento de identidad para firmar el consentimiento. Sin embargo, muchos estudiantes no pudieron poner el número de su documento de identidad y pusieron números aleatorios para llenar el espacio obligatorio y confirmar su aprobación para usar los datos. Tomando en cuenta esta realidad y los datos arrojados en este consentimiento informado, se usaron los siguientes datos para la autorización: número de celular, número de documento de identidad y el tiempo de enviar el formulario hecho.

¹¹ Refiere al apartado *Muestra y Población*

7. Resultados

En este apartado se presentan los resultados del análisis de las muestras escritas de los niveles A2, B1 y B2. En primer lugar, se abarca el análisis preliminar con el propósito de mostrar los resultados obtenidos en los diferentes paradigmas lingüísticos con la taxonomía tomada de Pico (2022). En segundo lugar, después de la elección del paradigma con la máxima frecuencia de errores (en este caso el paradigma nominal), se ahonda en cada nivel (A2, B1 y B2) para mostrar las frecuencias de los errores, primero desde el criterio lingüístico y luego desde el criterio etiológico. Todo este apartado de los resultados expone las frecuencias de los errores y luego se indaga sobre los rasgos comunes en cada nivel y los que atraviesan todos los niveles. En este orden de ideas, después de etiquetar todas las muestras escritas de los niveles A2, B1 y B2 en la plataforma *UAM corpus tool*, se encontró lo siguiente:

Tabla 40: *Análisis preliminar de errores de nivel A2 con 19 muestras*

Paradigma	Número	Porcentaje
Sistema nominal	46	61.33%
Sistema preposicional	11	14.67%
Sistema verbal	11	14.67%
Sistema pronominal	5	6.67%
Otros	2	2.67%
Total	75	

Nota. Elaboración propia

Tabla 41: *Análisis preliminar de errores de nivel B1 con 14 muestras*

Paradigma	N	Porcentaje
Sistema nominal	28	51.85%
Sistema preposicional	13	24.07%
Sistema verbal	7	12.96%
Sistema pronominal	6	11.11%
Otros	0	0.00%
Total	54	

Nota. Elaboración propia

Tabla 42: *Análisis preliminar de errores de nivel B2 con 27 muestras*

Paradigma	N	Porcentaje
Sistema nominal	61	45.52%
Sistema preposicional	26	19.40%
Sistema verbal	21	15.67%
Sistema pronominal	17	12.69%
Otros	9	6.72%
Total	134	

Nota. Elaboración propia

En nivel A2 se encontró que el sistema nominal tiene la máxima frecuencia de errores (46 errores) que corresponde a 61.33% de los errores en total. En nivel B1 también, el sistema nominal tiene alta frecuencia (28 errores) que corresponde a 51.85% del total. En nivel B2, también tiene la frecuencia más alta con 61 errores, que corresponde a 45.52% del total. Se puede observar, en los tres niveles, casi la mitad de los errores (más de 45%), sigue siendo el paradigma nominal, razón por la cual se decidió investigar en profundidad este paradigma y analizar los errores desde los criterios lingüístico y etiológico.

Por otro lado, los errores más frecuentes, en todos los niveles, después del paradigma nominal fueron en el paradigma preposicional, el verbal y por último el pronominal. En cuanto a los errores en el sistema preposicional, los estudiantes tenían dificultad en entender el significado de varias preposiciones. Por ejemplo: ‘*Nosotros fuimos en un restaurante tapas*’. En esta oración, el estudiante no tuvo claro el significado “dentro de” que ofrece la preposición ‘en’. Quizás, este error se dio por el menor abordaje de las preposiciones en la clase. En otro orden de ideas, los errores en el sistema verbal tenían otros tipos como la omisión, la adición y la concordancia entre otros. Por ejemplo: ‘*Estaba muy interesantes los pequeños*’. En conclusión, hay muchos otros ejemplos que se puede

analizar en el corpus CAELE/2, pero como el alcance de este estudio es analizar el paradigma con máximo número de errores, se abordó solo el paradigma nominal.

Para realizar el análisis de este sistema en todas las muestras de los diferentes niveles, A2, B1 y B2, se usó la taxonomía adaptada de Ferreira y Elejalde (2020), que se instaló en la plataforma *UAM corpus tool*, como ya fue mencionado en el apartado [7.2.2](#) del marco metodológico. A continuación, se expone el análisis y las frecuencias de los errores, primero desde el criterio lingüístico: la categoría gramatical, el aspecto sintáctico, la morfología y luego, desde el criterio etiológico: errores intralingüísticos e interlingüísticos.

7.1 Nivel A2

7.1.1 Criterio lingüístico

En el nivel A2 se encontró que hubo en total de 71 errores en el paradigma nominal desde el criterio lingüístico. Como se observa en la *figura 6*, el mayor número de errores se encontró en la elección de la categoría gramatical con 42 errores (59,91%), seguido del aspecto sintáctico con 20 errores (28,16%) y al final la morfología con 9 errores (12,67%). Es de destacar que en todos los niveles, la categoría gramatical y los aspectos sintácticos se explican desde nivel oracional y la morfología solo desde nivel palabra y no desde su interacción con la oración.



Figura 5: Errores lingüísticos en nivel A2

Ahora bien, en la categoría gramatical, se encontró que el artículo definido tenía el máximo número de errores: 29 que corresponde a 69,04% de total en esta categoría. El siguiente era el artículo indefinido con 5 errores (11,90%), el sustantivo con 4 errores (9,52%), el adjetivo con 3 errores (7,14%) y al final el determinante numeral con solo 1 error (0,23%). Como se puede percatar en las cifras y también en la *figura 7*, el artículo definido presentó la mayor dificultad en este nivel con otras categorías con pocos errores, por ejemplo, el determinante numeral, el sustantivo y el adjetivo.

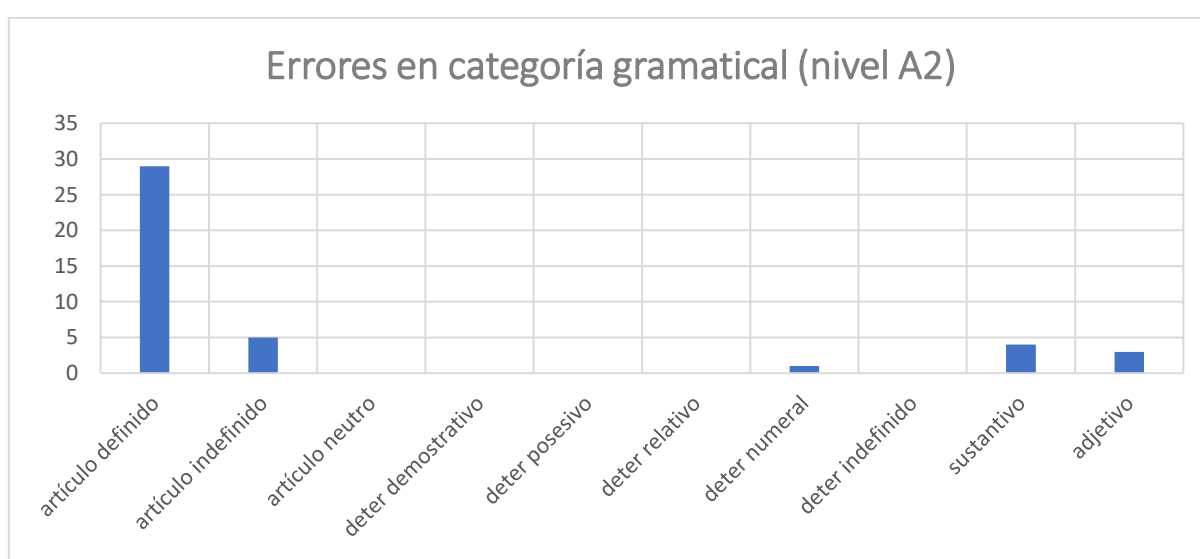


Figura 6: Errores en las categorías gramaticales de nivel A2

Sobre la base de las cifras anteriores, primero se exponen las categorías con poca frecuencia en los ejemplos mostrados en la *tabla 43* y luego, las categorías con alta frecuencia de errores. Cabe aclarar que no había ningún error en las siguientes categorías: artículo neutro, determinante demostrativo y determinante relativo.

Tabla 43: *Muestras de los errores en la categoría gramatical en nivel A2*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0276	actualmente estoy en	tres	años de bcom.aparte de esto,tambien estoy estudiando un diploma	Fs det num
ELEME0265	viaje internacional por	vocación	ya que se acerca el año nuevo?	Fs sust

ELEME0276	soy una persona religious que cree firmemente en	dias	.me encanta viajar una	Fs sust
ELEME0213	Las habitaciones estan grandes y muy	comida	. Voy a recomendar	Sus adj

Nota. Elaboración propia

En la muestra ELEME0276, el participante usó ‘tres’ en lugar de ‘tercer’. En las muestras ELEME0265 y ELEME0276 sucede algo similar, se usó ‘vocación’ en lugar de ‘vacaciones’ y ‘dias’ en lugar de ‘dios’. En los dos casos, el participante usó sustantivos con significados diferentes, pero con una morfología cercana.

En cuanto a los errores más frecuentes, el artículo definido (29 errores) representa más dificultad en comparación con el artículo indefinido (5 errores) y el artículo neutro (0 errores). Durante el análisis, la cifra más alta demuestra la omisión del artículo definido con 24 errores (70,58%), la adición con solo 3 errores (8,82%) y la falsa selección con 2 errores (5,88%). Solo hubo 3 casos de omisión del artículo indefinido (5,88%) y 2 casos de adición de él (8,82%). No se identificó ningún otro tipo de error del criterio descriptivo en el uso de los artículos. Basándonos en las cifras expuestas en la *figura 8*, se puede decir que la mayoría de los participantes cometieron errores de la omisión del artículo definido en nivel A2.

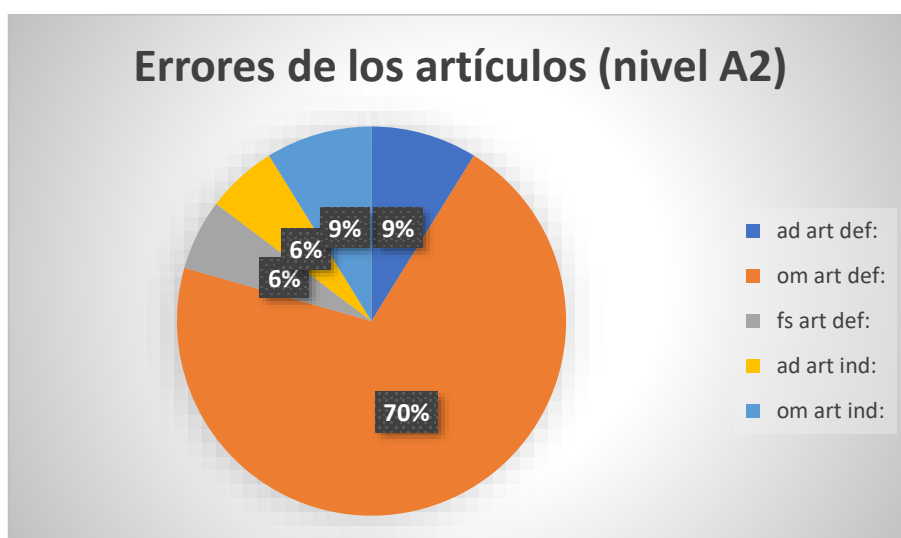


Figura 7: Errores de los artículos en nivel A2

Para ejemplificar las dificultades en los artículos se exponen algunos errores de los fragmentos de las muestras escritas de los participantes. Así, se evidencia que los participantes hindi hablantes tienen una confusión en los artículos y suelen omitirlos por razones interlingüísticas que abordamos en los siguientes apartados.

Tabla 44: *Muestras de los errores en los artículos en nivel A2*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME021 2	Me gusta encontrar	las	nuevas personas.	Ad_art_def
ELEME020 3	el cierre de		escuelas, instrucción remota	Om_art_def
ELEME020 3	el cierre de escuelas, instrucción remota en todo el país,		protestas por la justicia racial y elecciones, el papel de la educación nunca ha sido más crítico	Om_art_def
ELEME021 1	la tardes despues de		conferencia nosotros fuimos en	Om_art_def
ELEME027 6	dias.me encanta viajar	un	a nuevos lugares con	Ad_art_ind
ELEME027 5	He vivido en Nueva delhi.Nueva delhi es		muy hermosas ciudad.es la capital de la india.	Om_art_ind

Nota. Elaboración propia

La siguiente categoría es el aspecto sintáctico de la oración, en el cual hay dos subcategorías: el orden sintáctico y la concordancia. En el aspecto sintáctico en total hay 20 errores que corresponden a 28,16% de los errores lingüísticos. Como se ve en la *figura 9*, la mitad de los errores es de concordancia de género (10 errores), de concordancia de número son 6 errores y 4 errores del orden sintáctico.

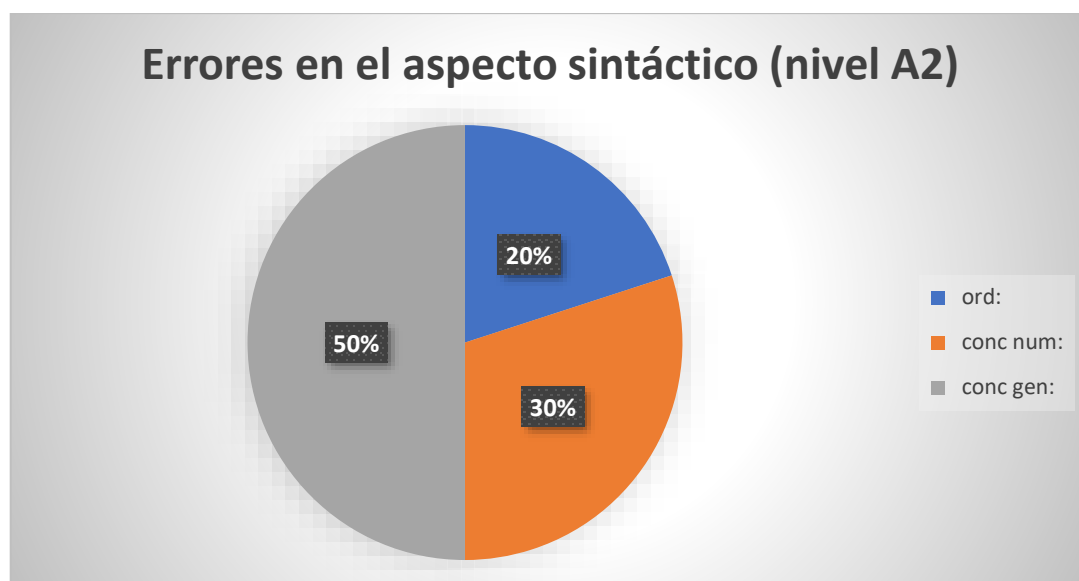


Figura 8: Errores en el aspecto sintáctico de nivel A2

En algunas muestras en específico como en ELEME0211 y ELEME0213, se presentaron más errores en la concordancia de género. Para mostrar mejor el panorama, se exponen en la siguiente *tabla 45*, estos errores comunes desde las muestras de los participantes.

Tabla 45: *Muestras de los errores en el aspecto sintáctico en nivel A2*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0211	Hay	muchas	edificios de architecte Antonio	Conc_gen
ELEME0252	El servicio de comida en el hotel no está muy	variada.		Conc_gen
ELEME0275	es mi lugar	favorita	porque es mi lugar	Conc_gen
ELEME0213	sirven la comida	vegetariano	. Hay una Grande	Conc_gen
ELEME0214	Los estudiantes hacen	los	actividades practicales. En	Conc_gen
ELEME0211	Por la	tardes	despues de conferencia nosotros	Conc_num
ELEME0275	los alimentos más	famous	de delhi.	Conc_num
ELEME0276	gente pero no demasiado	mi	aficiones es bailar en	Conc_num

ELEME02	Es una	Grande	calle que va hasta el	Ord
11				

Nota. Elaboración propia

Se observa que todas las muestras de arriba tienen errores en la concordancia de género y número menos el último que representa un error en el orden sintáctico. Después de ver que hay casos y evidencias de intentos de concordar el sustantivo con adjetivos y artículos, se llega a la conclusión de que los participantes podrían tener dificultades en la elección del sustantivo con lo cual tenían que concordar el adjetivo o quizás no sabían los respectivos géneros de los sustantivos. En cuanto al orden sintáctico, hay evidencia clara del uso de adjetivo antes del sustantivo y una confusión en la creación de la frase respetando el orden sintáctico y su significado correspondiente. El mismo error se puede entender desde la falta de conocimiento de la apócope del adjetivo *grande*, porque la misma oración se puede corregir así: *es una gran calle*. Este tema se va a explorar en el criterio etiológico para entender la raíz de este error.

Tabla 46: Muestras de los errores de la morfología en nivel A2

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0211	estuve la casa	batallo	qui es rosa.	Fs morf
ELEME0213	Hay tres	bars	y dos restaurants muy	Fs morf
ELEME0213	niños. Las	habitacions	están grandes y muy	Fs morf
ELEME0275	están llenas de	peatons	por todas partes se	Fs morf
ELEME0275	tiene un clima	extremist	hace mucho calor en	Fs morf
ELEME0276	estoy en	tres	años de bcom.aparte de esto	Fs morf

Nota. Elaboración propia

En términos de morfología a nivel palabra, hubo 9 casos con una tendencia general de errores en la pluralización de los sustantivos. En la muestra ELEME0211, el participante usó ‘batallo’ en lugar de ‘batalla’ o un nombre propio. En las muestras ELEME0213 y ELEME0275, se escribió ‘bars’ en lugar de ‘bares’, ‘habitacions’ en lugar de ‘habitaciones’, ‘peatons’ en lugar de ‘peatones’ y ‘extremist’ en lugar de ‘extremista’.

Después de analizar estas muestras, es probable que los participantes se confundieron entre la morfología de inglés y español porque usan la regla de la gramática del inglés para pluralizar los sustantivos en español. No obstante, en la muestra ELEME0276, se usó ‘tres’ en lugar de ‘tercer’. Es probable que el participante se haya confundido en la morfología de los números cardinales y ordinales.

7.1.2 Criterio etiológico

Se analizaron los errores desde la transferencia de la lengua materna- hindi y la confusión propia de la lengua meta, que en este caso era el español. Así, desde el aspecto intralingüístico, se indagaron los errores de simplificación, sobregeneralización e hipercorrección. Como resultado de este análisis, se expone la *figura 10*. Casi el 75% de los errores son interlingüísticos (38 errores) y los restantes son errores intralingüísticos (13 errores). Esto quiere decir que los estudiantes de nivel A2 tienen la interlengua más permeada de la lengua materna hindi que la lengua meta español.

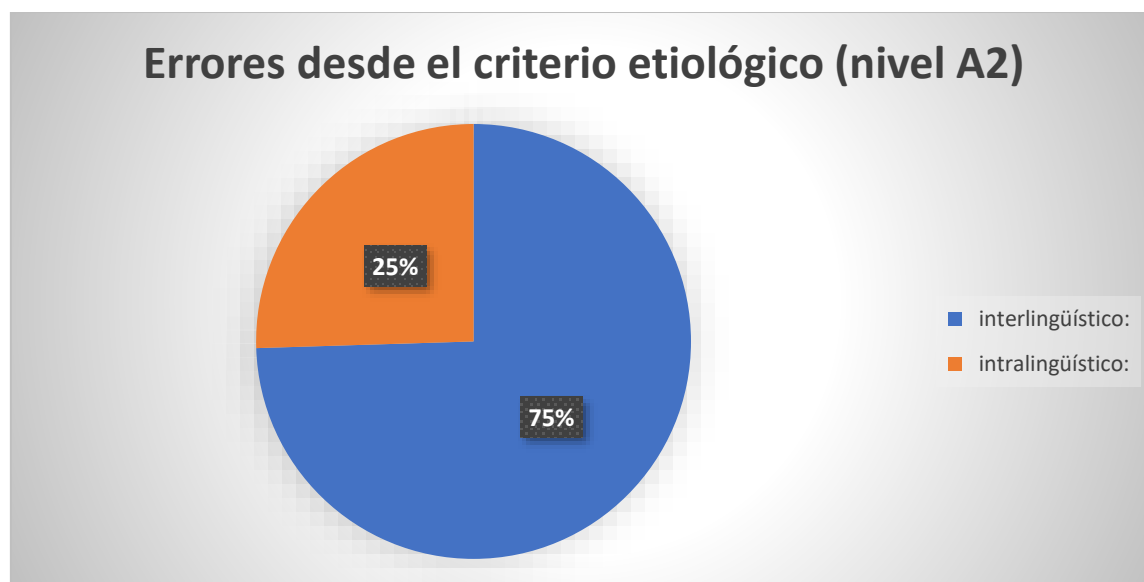


Figura 9: Errores desde el criterio etiológico en nivel A2

Ahora, para entender mejor los errores desde las dos subcategorías, se exponen algunas muestras y se discuten algunos errores en concreto.

Tabla 47: *Muestras de los errores desde el criterio etiológico (interlingüístico) en nivel A2*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0203	Después de un año que involucró una pandemia mundial, el cierre de		escuelas, instrucción remota en todo el país, protestas por la justicia racial y elecciones,	ETIO_LM
ELEME0211	elegante. Hay	muchas	edificios de architecte Antonio	ETIO_LM
ELEME0211	Hay muchas edificios de		architecte Antonio Gaudi.	ETIO_LM
ELEME0211	la tardes despues de		conferencia nosotros fuimos en	ETIO_LM
ELEME0275	es mi lugar	favorita	porque es mi lugar	ETIO_LM
ELEME0276	años.estoy estudiando bcom de		universidad de delhi y	ETIO_LM
ELEME0276	dr español de	hansraj colegio	.soy una persona ambivertida	ETIO_LM
ELEME0211	rosa. Hay	muchas	taxis en Barcelona y	ETIO_LM
ELEME0211	. Es una	Grande calle	que va hasta el	ETIO_LM
ELEME0211	que va hasta	el	playa. Hay muchos	ETIO_LM

Nota. Elaboración propia

El primer error más común en la transferencia de hindi es la omisión de artículo. Este error ya se ha destacado como el error más común en las muestras de nivel A2 desde el criterio lingüístico y también a partir de la explicación en el análisis contrastivo entre hindi y español en el marco teórico. En este análisis se destacó la ausencia de los artículos en hindi y la importancia de ellos en español, que resulta en uno de los desafíos principales en el aprendizaje de español para hindi hablantes. El otro error común es la concordancia de género y número (ELEME0211, ELEME0275). En primer lugar, según la *tabla 47*, los participantes tienen dificultades en identificar el género del sustantivo y lo asignan según el sustantivo en hindi. Abajo se exponen algunos ejemplos para entender esta transferencia (*tabla 48*).

Tabla 48: *Muestras de cambio de género de español a hindi*

Palabra en español	Género en español	Palabra en hindi	Género en hindi
edificio	masculino	imarat इमारत	femenino
lugar	masculino	Jagah जगह	femenino
playa	femenino	Samudra tat समुद्र तट (o palabra prestada de inglés-beach)	masculino
taxi	masculino	Gaddi गाड़ी (o palabra prestada de inglés- cab)	femenino

Nota. Elaboración propia

En segundo lugar, en hindi la concordancia se refleja tanto en el adjetivo (con excepciones en el caso de los adjetivos sin flexión y los adjetivos con flexión, pero con declinaciones) como en la conjugación del verbo (sistema pronominal)¹², en cambio en español es solo en el adjetivo. Por ejemplo: *Badi sadak*, *bada shēhēr* pero *badi sadakēin* (como se ha visto en el plural femenino no cambia el adjetivo¹³) Por eso, en las muestras se ejemplifican algunos errores de concordancia de género y número, aunque los estudiantes sepan el concepto de la concordancia en el paradigma nominal en hindi. Es de destacar que los estudiantes hindi hablantes conocen las reglas gramaticales de hindi inconscientemente por ser su lengua materna y necesitan un abordaje explícito en el aula para poder explotar este conocimiento lingüístico.

El último error, no con mucha frecuencia, es el del orden sintáctico en las muestras ELEME0211 y ELEME0275. En la primera frase “Es una grande calle que va hasta el” tiene el error en orden sintáctico y la falta de apocopar el adjetivo antepuesto: ‘gran calle’ en lugar de ‘grande calle’. La posición del adjetivo y su correspondiente significado representa una confusión en el posicionamiento del adjetivo en español. Por ejemplo, cuando ‘grande’ va antes del sustantivo, se convierte en ‘gran’ en ciertos casos como ‘gran

¹² Ejemplo: Mai khana *khata* hun, mai khana *khati* hun. En este caso, en la primera frase, el pronombre yo es masculino y en el segundo es femenino.

¹³ Refiere al apartado sobre el contraste de adjetivos en español e hindi.

hombre, gran chica’ y al mismo tiempo cambia su significado original de representar el tamaño de algo o la edad de alguien. En el caso del ejemplo, el estudiante quería hablar del tamaño de la calle que puede hacer tanto con *gran* como con *grande* pero no apocopó el adjetivo pospuesto.

En cuanto a los errores intralingüísticos, la categoría con mayor número de errores fue la sobregeneralización con 7 errores (13,72%), 4 errores de simplificación (7,84%) y 1 error de hipercorrección (1,96%). Abajo se exponen algunos ejemplos concretos con el objetivo de contextualizar y entender el porqué de los errores.

Tabla 49: *Muestras de los errores desde el criterio etiológico (interlingüístico) en nivel A2*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0211	Por la	tardes	despues de conferencia nosotros	ETIO_SOBR E
ELEME0211	por que hay	muchas	gente qui van en	ETIO_SOBR E
ELEME0213	sirven la comida	vegetariano	. Hay una Grande	ETIO_SOBR E
ELEME0213	niños. Las	habitacions	están grandes y muy	ETIO_SOBR E
ELEME0205	El problema que ocurre con las redes sociales es que	las redes sociales	brindan tantas cosas que a veces las personas simplemente creen	ETIO_SIM
ELEME0276	actualmente estoy en	tres	años de bcom.aparte de	ETIO_SIM
ELEME0276	estoy en tres	años	de bcom.aparte de esto	ETIO_SIM
ELEME0212	Me gusta encontrar	las	nuevas personas.	ETIO_HIPER

Nota. Elaboración propia

En esta *tabla*, es evidente que la sobregeneralización demuestra una tendencia de errores de concordancia, tanto de género como de número. Los estudiantes suelen usar las mismas reglas que aprendieron anteriormente para nuevos elementos. Para ilustrar, en nivel básico, un estudiante de español aprende a decir primero ‘buenas tardes’ y así el participante lo extrapoló a decir ‘la tardes’ en vez de ‘la tarde’ porque podría pensar que

quizás así existe la palabra morfológicamente. Lo mismo sucede con el ‘ser vegetariano’ y ‘comida vegetariana’: el estudiante podría pensar que la palabra solo existe en género masculino morfológicamente, ‘habitacions’ con una regla generalizada de cómo pluralizar un sustantivo. Estos ejemplos respaldan la tendencia general de usar siempre la flexión masculina y singular en niveles básicos. Más adelante, en nivel B1 y B2 también hay ejemplos del uso de adjetivos en masculino con estudiantes femeninas.

En el caso de la simplificación, muchas veces los estudiantes, por tener miedo a cometer errores o no tener suficientes herramientas o no saber o tener dudas en el uso de algún elemento gramatical, suelen simplificar la oración. Por ejemplo, la repetición de ‘las redes sociales’ en ELEME0205 mostraría una redundancia que le ayudaría a conjugar el siguiente verbo. De la misma forma, el uso de ‘tres’ es una forma de evitar usar el número ordinal- ‘tercero’. Al final, en la muestra ELEME0212, la adición del artículo definido representa una hipercorrección porque usó un elemento gramatical donde no se debería usar. Esta hipercorrección puede ser un resultado de sobre corrección de un error en el aula o simplemente un intento de usar un artículo siempre con un sustantivo.

7.2 Nivel B1

7.2.1 Criterio lingüístico

En el nivel B1, se encontró que había 60 errores en total. Como se puede ver en la *figura 11*, el mayor número de errores está en la elección de la categoría gramatical con 46 errores (76,66%), seguido del aspecto sintáctico con 10 errores (16,66%) y al final la morfología con 4 errores (6,66%).

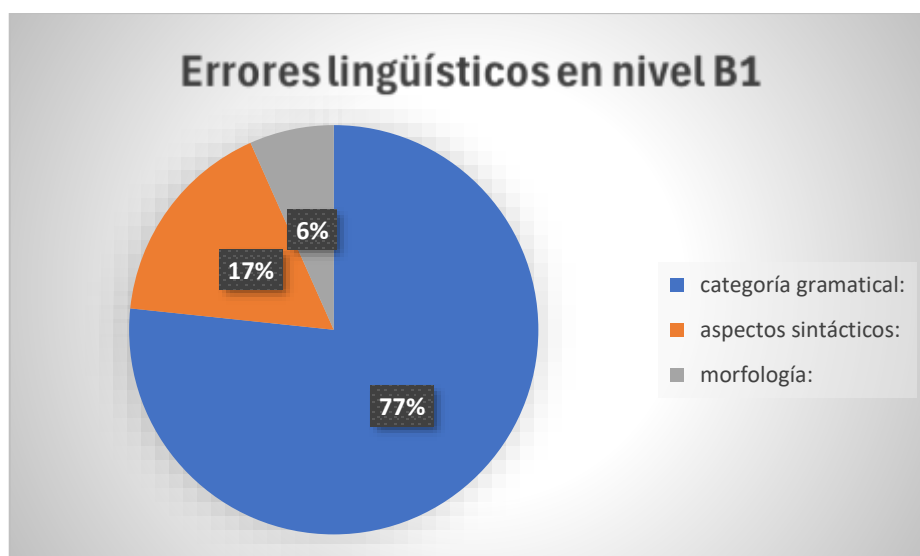


Figura 10: Errores lingüísticos en nivel B1

En la categoría gramatical, el mayor número de errores se encontró en el artículo definido con 22 errores (47, 82%), los siguientes, el artículo indefinido y el sustantivo con 8 errores cada uno (12,50%), luego, determinante posesivo con 6 errores (9,38%) y determinante demostrativo y adjetivo con solo 1 error cada uno (2,10%). En este nivel, artículo neutro, determinante relativo, determinante numeral y determinante indefinido no presentaron ningún error. Estas cifras junto con la representación gráfica en la *figura 12* evidencian los errores más frecuentes y menos frecuentes. En el mismo orden de ideas, primero se exponen los fragmentos de errores con menos frecuencia como el determinante demostrativo, el adjetivo y el determinante posesivo y después, se abordan las categorías con errores frecuentes que en este nivel siguen siendo primero, el artículo definido y luego el artículo indefinido.

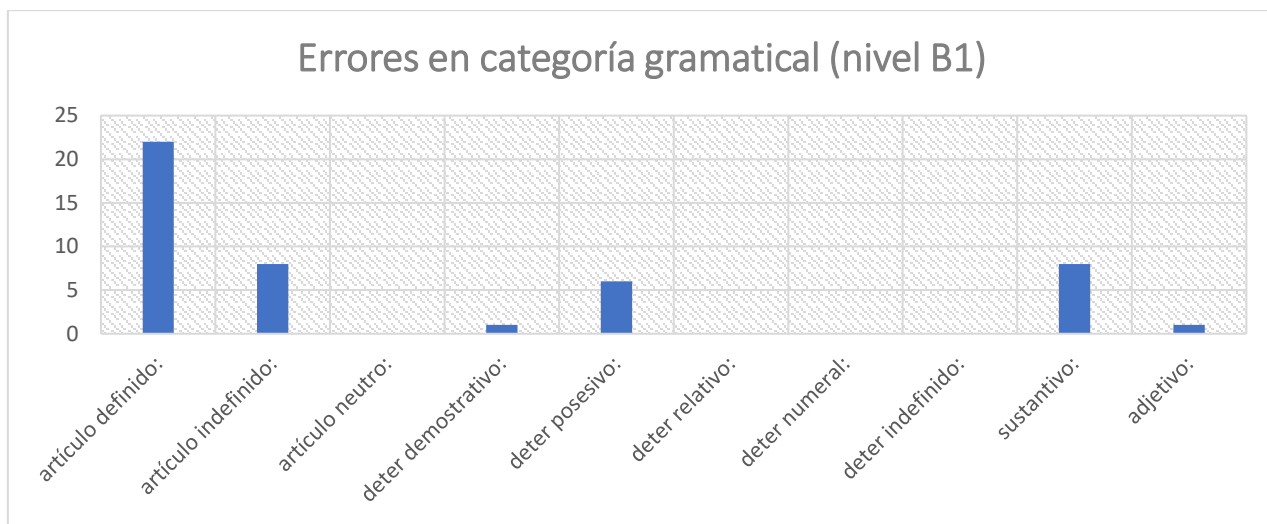


Figura 11: Errores en categoría gramatical de nivel B2

Tabla 50: Muestras de los errores menos frecuentes en nivel B1

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0270	Mi nombre es Urvashi.	Me	apellido es Varshney	Sus det pos
ELEME0270		Me	Cumpleaños es el vientesiete	Sus det pos
ELEME0270		Me	correo electrónico es	Sus det pos
ELEME0270	Mi padre es un	hambre	De negocios.	Fs sust
ELEME273	Había muchas	transportaciones	como autobuses	Fs sust
ELEME273	La ciudad tiene más	populación	Y mucha contaminación	Fs_sust

Nota. Elaboración propia

Como se puede observar, el estudiante de la muestra ELEME0270, en lugar de usar el determinante posesivo ‘mi’, usó una categoría gramatical completamente diferente: el pronombre personal reflexivo de primera persona singular. Es de destacar que, en esta muestra, la confusión empezó por la conjugación de los verbos pronominales como es el caso de ‘apellidarse’ y luego, se presentó en otras categorías gramaticales también. Para terminar este apartado, se señala la falsa selección de los sustantivos que, principalmente, surgió por los falsos cognados traídos del idioma inglés, con la excepción al error de sustantivo ‘hambre’. Este error en particular, aunque difícil de razonar, puede ser una confusión entre las vocales ‘o’ y ‘a’.

En cuanto a los errores más frecuentes en la categoría gramatical, como se ha mencionado antes, los artículos tuvieron mayor frecuencia. En esta categoría, se presentaron errores de omisión (19 errores, 63,33%) y de sustitución (3 errores, 10%) del artículo definido y de omisión (7 errores, 23,33%) y de falsa selección (1 error, 3,33%) en el caso del artículo indefinido.

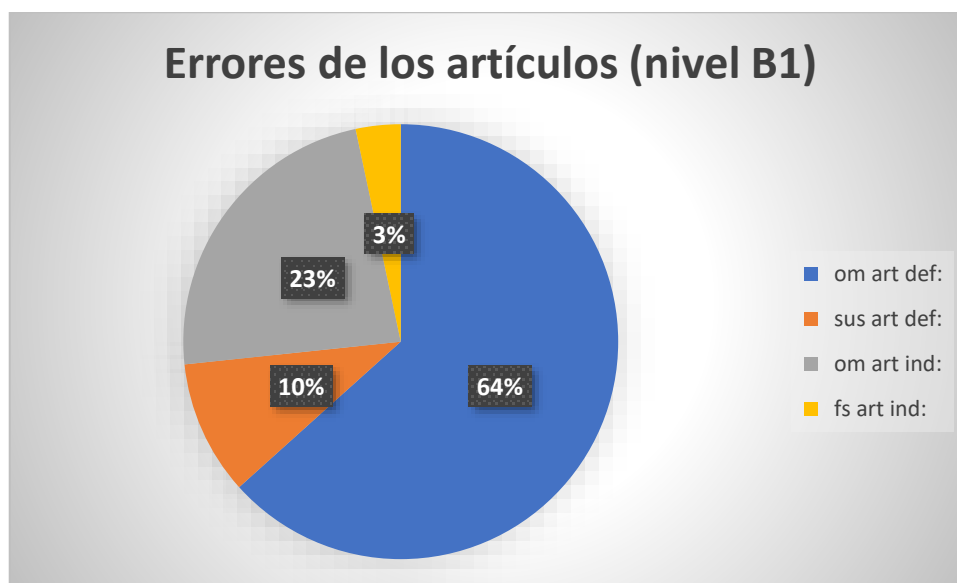


Figura 12: Errores de los artículos en nivel B1

A continuación, se exponen algunos fragmentos de las muestras para observar y entender el contexto de los errores cometidos.

Tabla 51: *Muestras de los errores en los artículos*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0270	Me gusta come comida	de	sur de India	Om art def
ELEME0273	También he visitado		Parque de nacional Zoologicál	Om art def
ELEME0273	También he visitado Parque de nacional Zoologicál y		Museo Nacional	Om art def
ELEME0274	Soy		Más Joven de mi familia.	Om art def
ELEME0274	Estoy estudiando		Diploma en español en colegio Hansraj	Om art ind
ELEME0281	Siempre hay un problema de		Red, preste atención	Om art def

ELEME0281	Brinda un buen servicio a	tus	clientes	Fs det pos (El texto escrito en usted)
-----------	---------------------------	-----	----------	---

Nota. Elaboración propia

Los errores de la omisión de los artículos, tanto definido como indefinido, siguen siendo el error más común, como en nivel A2. No obstante, en nivel B1, se encuentran errores en la formación de oraciones más complejas como ‘Soy la más joven de la familia’. Aquí el uso de artículo no es un uso común de un sustantivo, sino como un elemento de comparación en el superlativo.

Al final, algo que se encontró en este nivel con frecuencia fue el cambio de la forma de tratamiento ‘usted’ al principio y ‘tú’ al final. Por ejemplo, en el caso de la muestra ELEME0281, el estudiante usó ‘tus clientes’ al final del texto cuando casi todas las frases estaban escritas en *usted* como la forma de tratamiento al principio. Este error puede que sea una dificultad o un esfuerzo extra de mantener la conversación con la forma de tratamiento usted.

En el aspecto sintáctico de la oración, hubo algunos casos de errores en la concordancia y el orden sintáctico. Abajo, en la *figura 14* se puede ver el porcentaje de los errores, con el más alto: la concordancia de número con 5 errores (50 %), el orden sintáctico con 3 errores (30 %) y al final, 2 errores de concordancia de género (20%).

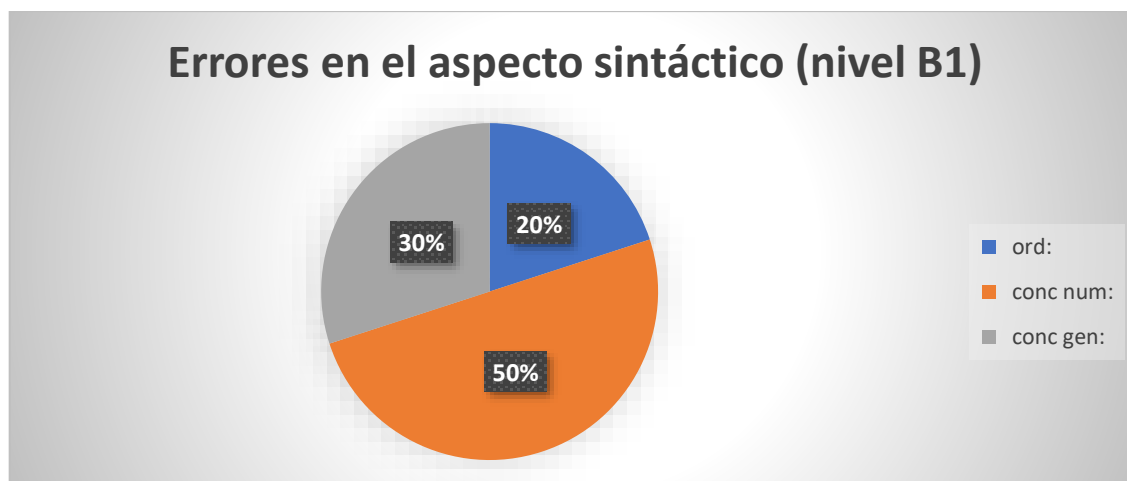


Figura 13: Errores sintácticos en nivel B1

Con el fin de entender mejor los errores en el aspecto sintáctico, se exponen algunos fragmentos. En los errores de orden sintáctico, también como hemos visto en nivel A2, los estudiantes suelen usar los adjetivos antes de los sustantivos (ELEME0271). En cuanto a los errores de la concordancia de número, la mayoría de los errores son de concordancia entre sustantivo y adjetivo. Por ejemplo, ‘estudiantes’ no concuerda con ‘rapido’. De la misma forma, los errores en la concordancia de género demuestran el mismo patrón. Por ejemplo, ‘la gente’ no concuerda con ‘este’, ‘museo’ con ‘esta’ y ‘lugar’ con ‘buena’ (Tabla 52).

Tabla 52: Muestras de los errores en el aspecto sintáctico de nivel B1

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0271	mucho. Hay	muchos famosos lugares	a visitar allí y	ord
ELEME0271	allí son muy	amable	porque ellos aydaronme mucho	Conc num
ELEME0272	y estoy usando	sus	telefonía celular desde hace	Conc num
ELEME0273	arroz. Había	mucho	monumentos antigous y famosos	Conc num
ELEME0274	. Ambos son	mayor	de mí. Soy	Conc num
ELEME0278	momento estudiantes son	rapido	en antes timpo Según.el	Con num
ELEME0271	la gente como	este	también. Y la	Conc gen
ELEME0273	museo Nacional.	Esta	lugar fue muy Buena	Conc gen
ELEME0273	lugar fue muy	Buena	. Prefiero que visite	Conc gen

Nota. Elaboración propia

En cuanto a la morfología a nivel palabra, hubo 4 errores en este nivel. La tendencia general es la confusión en las palabras pluralizadas. En la tabla 52 se exponen estos casos específicos, por ejemplo, ‘pantalons’ en lugar de ‘pantalones’, ‘muchas’ en lugar de ‘mucho’, ‘aficciones’ en lugar de ‘aficiones’. El último ejemplo representa otra dificultad

morfosintáctica de no poder elegir la forma adecuada (plural o singular), por eso el participante escribió, ‘antes timpo’ en lugar de ‘tiempos anteriores’ o ‘anteriores tiempos’.

Tabla 53: *Muestras de los errores de la morfología en nivel B1*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0270	llevo camisieta y	pantalons	. Me gusta comer	Fs morf
ELEME0273	simpática . Había	muchas	transportaciones como autobuses ,	Fs morf
ELEME0274	Delhi. Mis	aficciones	es viajar , leer	Fs morf
ELEME0278	son rapido en	antes timpo	Según.el trabajo inteligente es	Fs morf

Elaboración propia

7.2.2 Criterio etiológico

Desde el criterio etiológico, en nivel B1, se observa en la figura 15, que el 66,66% (30 errores) de los errores son de naturaleza interlingüística, mientras que el 33,33% restante son de tipo intralingüístico. Dentro de los errores intralingüísticos, 10 errores son de simplificación (22,22%), 3 errores de sobregeneralización (6,66%) y 2 errores de hipercorrección (4,44%).



Figura 14: Errores desde el criterio etiológico en nivel B1

Abajo primero se expondrán los errores desde la interferencia de hindi y luego los propios del proceso del aprendizaje de español.

Tabla 54: *Muestras de los errores de la interferencia de hindi en nivel B1*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0270	Me gusta come comida	de	sur de India	ETIO LM
ELEME0273	También he visitado		Parque de nacional Zoologicál	ETIO LM
	También he visitado Parque de nacional Zoologicál y		Museo Nacional	ETIO LM
ELEME0271	mucho. Hay	muchos famosos lugares	a visitar allí y	ETIO LM
ELEME0278	Estudiantes son rapido en	antes timpo	Según.el trabajo inteligente es	ETIO LM
ELEME0273		Esta	Lugar fue muy buena.	ETIO LM
ELEME0273	Esta lugar fue muy	buena		ETIO LM
ELEME0274	Estoy haciendo pos	graduación	De MDU	ETIO LM

Nota. Elaboración propia

Para empezar, casi todos los errores de omisión de los artículos son provenientes del hindi, dada la ausencia de este componente gramatical. En el caso del orden sintáctico, la muestra ELEME0271, ‘muchos famosos lugares’ sigue el mismo orden sintáctico que en hindi ‘kaee prasiddh jagahein, कई प्रसिद्ध जगहें’. De la misma forma, en la muestra ELEME0278, el error ‘antes tiempo’ representa el orden sintáctico de hindi- ‘पहले का समय, *pahalē kā samaya*’.

En la muestra ELEME0273, el estudiante consideró que el sustantivo ‘lugar’ era femenino y por la tanto, concordó el adjetivo con ‘buena’. Como ya se había explicado en los resultados de nivel A2, el sustantivo ‘lugar’ en hindi es femenino y en español, masculino. Este error se repitió en este nivel también. Al final, la palabra ‘graduación’ se refiere a ‘graduation’ de inglés. Cabe señalar que es una palabra prestada de inglés y

normalizada en el uso diario en India, especialmente por los hindi hablantes. Por eso, este error de falso cognado se dio por la interferencia de la lengua materna.

Ahora bien, este último apartado se dedica a explicar los errores desde el criterio intralingüístico. Abajo, se exponen algunos fragmentos para contextualizar los errores y entenderlos desde el criterio etiológico.

Tabla 55: *Muestras de los errores de la interferencia de español en nivel B1*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0270	es Urvashi.	Me	Apellido es Varshney.	ETIO SOBRE
ELEME0271	la gente como	este	también. Y la	ETIO SOBRE
ELEME0274	. Ambos son	mayor	de mi. Soy	ETIO SIM
ELEME0218	Estoy nueva en esta ciudad y muy emocionada para explorar	La ciudad	Con un nativo	ETIO SIM
ELEME0274	Soy		Más joven de mi familia	ETIO SIM
ELEME0279	Vamos a sacar provecho de Internet en lugar de ser un	beneficio	para el internet	ETIO SIM
ELEME0281	Cliente y solucionar	su	Problema lo antes posible	ETIO HIPER

Nota. Elaboración propia

En el caso de las muestras ELEME0270, el estudiante sobregeneralizó el uso de ‘me’ a ‘mi’ y lo repitió en su texto. Muchas veces el error de sobregeneralización también es común desde la concordancia de número y género, como es evidente en el caso de la muestra ELEME0271 en la cual, se usa el demostrativo ‘este’ con ‘gente’ en vez de ‘esta’.

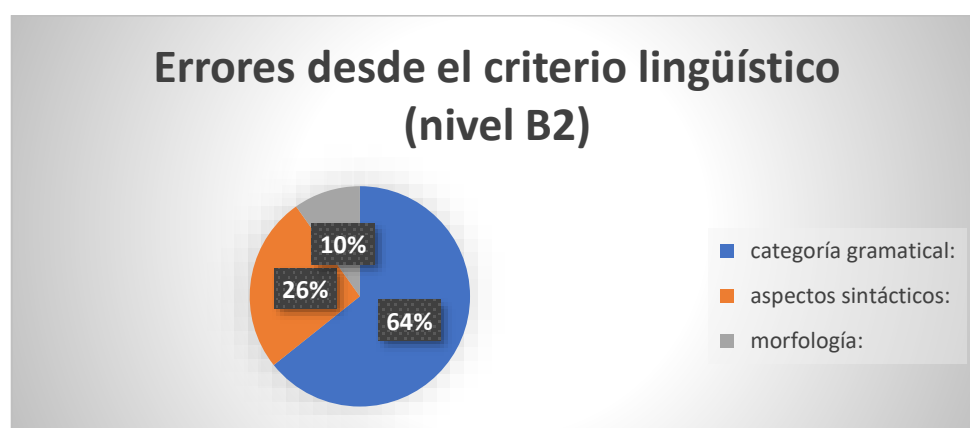
La simplificación, con máximo números de errores en el criterio intralingüístico, se evidencia en las muestras ELEME0218, ELEME0274 y ELEME0279 entre otras muestras escritas. Por ejemplo, la repetición de ‘la ciudad’ evidencia un intento de simplificar la frase para el entendimiento propio y así evitar recurrir a los recursos de nivel B1 como el

pronombre de objeto directo. En la frase, ‘soy ~~la~~ más joven’, la presencia del artículo ‘la’ presenta una dificultad de nivel intermedio porque este uso del artículo definido es para mostrar la superioridad¹⁴. El uso de la palabra ‘beneficio’ en lugar de ‘beneficiario’ también muestra una simplificación de léxico. Por último, en el nivel B1, solo se encontraron dos casos de hipercorrección y los dos con el uso de posesivos. En las dos oraciones el estudiante en lugar de usar los artículos usó los posesivos, lo que se explica desde la transferencia de hindi por la falta de los artículos en la lengua. Como ya se había dicho antes¹⁵, el posesivo juega un papel más importante en hindi por la ausencia completa de los artículos en la lengua, los aprendientes de español hindi hablantes suelen extrapolar esta regla y usar los posesivos como si fueran los artículos en español.

7.3 Nivel B2

7.3.1 Criterio lingüístico

En nivel B2, se encontraron 112 errores en 27 muestras en total. El mayor número de errores se encontró en la elección de la categoría gramatical con 72 errores que correspondía a 64,28% de los errores, seguido del aspecto sintáctico con 29 errores (25,89%) y al final la morfología con 11 errores que correspondía a 9,82% de los errores en total. En la *figura 16* se presentan estos porcentajes para visualizar los errores con sus diferentes frecuencias.



¹⁴ Un uso del artículo definido que se explica en niveles intermedios.

¹⁵ Refiere al aparatado de los posesivos en el contraste entre hindi y español.

Figura 15: Errores lingüísticos en nivel B2

A continuación, primero se expondrán los errores en la categoría gramatical por el alto porcentaje de errores y luego el aspecto sintáctico y la morfología. En este orden de ideas, en la *figura 17* se puede observar el error más común, con un alto porcentaje de errores en artículo definido con 63 errores, que corresponde a 87,50% de los errores en total. Las otras categorías tienen poca frecuencia de errores, por ejemplo: artículo neutro, determinante indefinido y adjetivo con solo 1 error (1,39%) cada uno. En cambio, determinante posesivo y sustantivo con 3 errores (4,17%) cada uno.

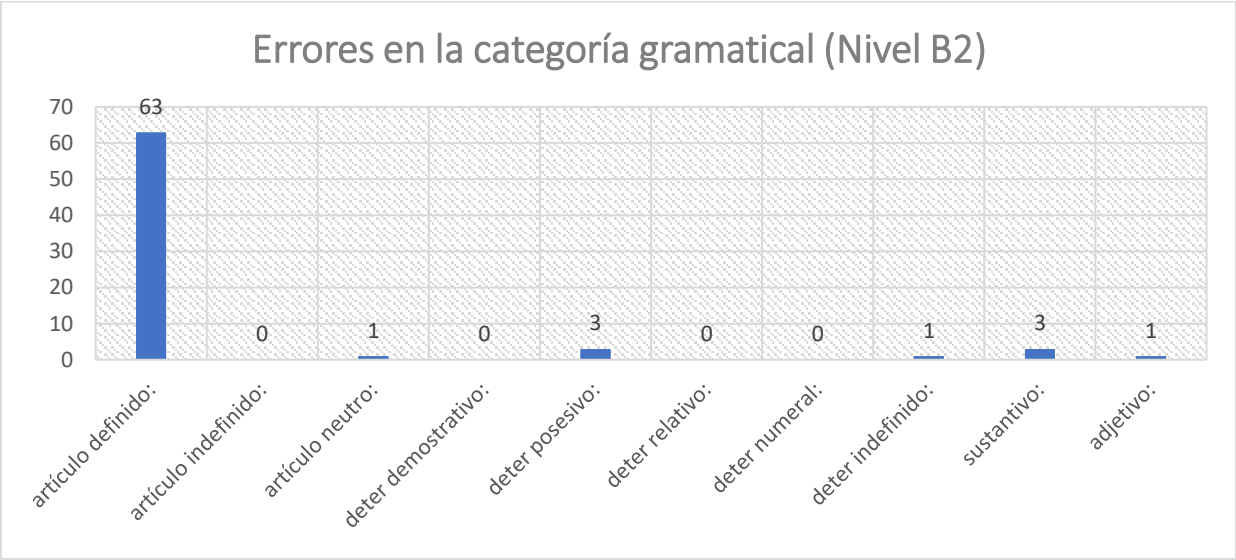


Figura 16: Errores en la categoría gramatical en nivel B2

Abajo se exponen primero las categorías con poca frecuencia y luego los artículos definidos con máximo número de errores.

Tabla 56: Muestras de los errores en la categoría gramatical en nivel B2

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0143	Es que la educación en el contexto de mi país,, no era la educación	más	importante como hoy pero a mi me gustaria describir	Sus adj

ELEME014 4	conectado con sus amigos		familia y el mundo	Om det pos
ELEME025 9	que es tranquilo.		Mamá manejar desayuno y	Om det pos
ELEME025 9	algunas tareas y ayudar		mamá con cosas alrededor	Om det pos
ELEME014 3	podemos usar durante	estudiar	. Personalmente , cuando	Sus sust
ELEME025 5	de mi primer	documentari o	independiente que viene,	Fs sust
ELEME026 2	Imagen en este	postre	de aplicación social usar	Sus sust
ELEME022 5	equipo hacen muy	bien	trabajo. Tenía la	Sus adj

Nota. Elaboración propia

En la muestra ELEME0143, el estudiante usó ‘más’ en vez de usar el adjetivo ‘tan’, por lo tanto, es la sustitución de una categoría gramatical, en este caso es un adverbio en lugar de un adjetivo comparativo. En el caso de los posesivos, los 3 errores se dieron por la omisión. En los 3 casos podría ser que, al participante, este elemento le pareció de baja importancia por su ausencia completa en su lengua materna y por eso se presentó la omisión de este elemento gramatical. En caso de los errores de los sustantivos, se puede decir que el participante usó el verbo ‘estudiar’ y los sustantivos ‘documentario’ y ‘postre’ en lugar de ‘los estudios’, ‘documental’ y ‘cartel’ respectivamente por una posible interferencia del inglés. Dado que, hay muchas palabras prestadas del inglés al hindi que se usan directamente sin traducir, por ejemplo: ‘postre’ y ‘documentario’ en este caso podría ser una interferencia de la palabra ‘poster’ y ‘documentary’. Al final, en el caso de la muestra ELEME0225, el estudiante escribió ‘bien trabajo’ en lugar de ‘buen trabajo’ por un error común que surge por la dificultad para diferenciar entre un adverbio y un adjetivo.

Ahora bien, se aborda la categoría de errores con alta frecuencia: los artículos definidos. Abajo se expone la figura 18 para visualizar el porcentaje de errores en esta categoría desde su descripción.

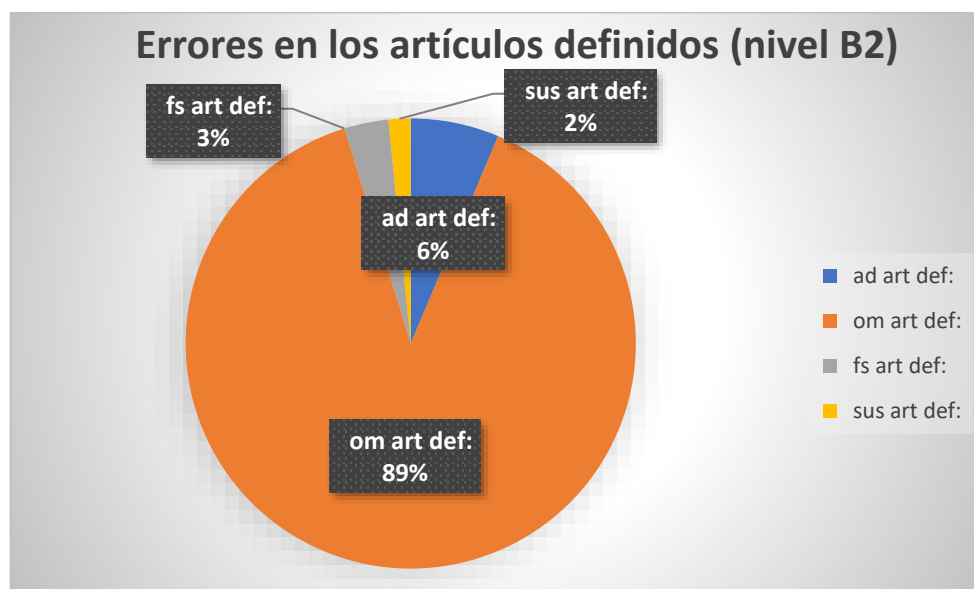


Figura 17: Errores en los artículos definidos en nivel B2

Se puede observar que la omisión es el error más común, lo mismo que en nivel A2 y B1, con el máximo número de errores: 56 errores (88,895), la adición con 4 errores (6,35%), la falsa selección con 2 errores (3,17%) y la sustitución con 1 error (1,59%). A continuación, se exponen las muestras de los fragmentos de los estudiantes para entender el contexto de los errores.

Tabla 57: Muestras de los errores en los artículos en nivel B2

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0040	de las actividades,		transportes, clima etc	Om art def
ELEME0210	la inteligencia artificial y		herramientas como ChatGPT,	Om art def
ELEME0223	La cultura,		comida , festival,	Om art def
ELEME0238	donde vende comida de		calle y es muy	Om art def
ELEME0228	liberada por likes y		comentarios refuerza esta conexión	Om art def
ELEME0145	los niños están muriendo de hambre no tienen	los	padres ni la comida ni casa.	Ad art def

ELEME0260	y bailar Y	famoso de las lugares	de la India ejemplo	Fs art def
ELEME0261	El teléfono de	mi	Compañía Apple, no	Sus art def

Nota. Elaboración propia

Se evidencia de que los errores de omisión de los artículos definidos son los más frecuentes (*tabla 57*). En todas las muestras presentes, se puede observar que a los participantes les costó entender el uso adecuado de los artículos y por eso la omisión de este elemento ocurre con tanta frecuencia. Lo mismo sucede con la adición del artículo definido, se agregan sin tener la necesidad de ello. En las últimas tres muestras, se usó ‘una’ en lugar de ‘la’ y ‘las lugares’ en lugar de ‘los lugares’. Al final, un error concurrente en todos los niveles también fue el uso de los posesivos en lugar de los artículos que se explica luego desde el criterio etiológico.

Desde el aspecto sintáctico, hubo 7 errores de orden sintáctico (24,13%), 7 errores en la concordancia del número (24,13%) y 15 errores en la concordancia de género (51,7%). Se puede visualizar estas cifras en la *figura 19*.

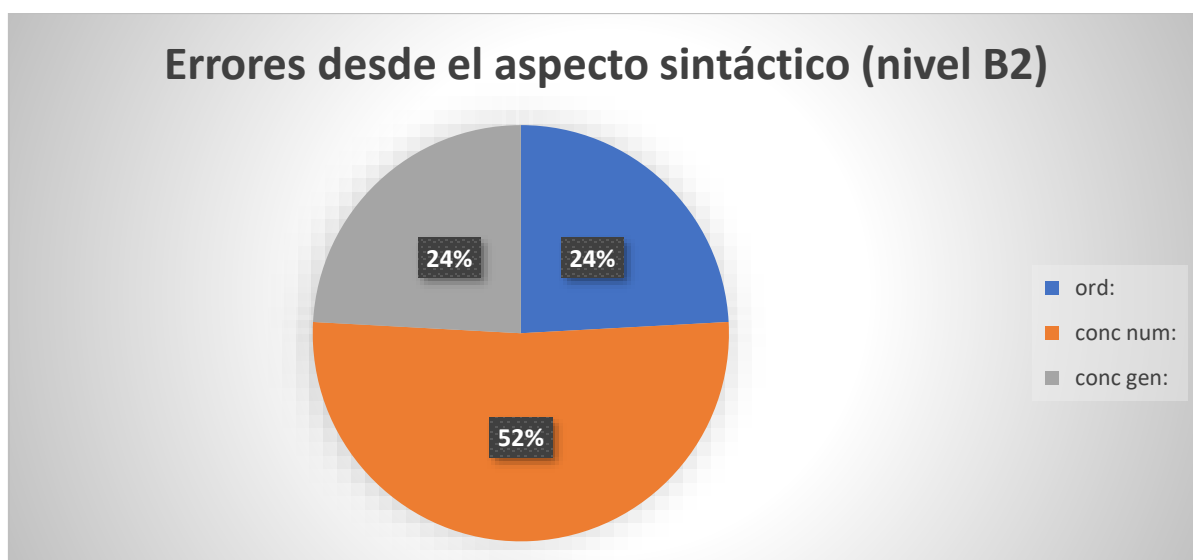


Figura 18: Errores desde el aspecto sintáctico en nivel B2

Para contextualizar los errores, se exponen algunos fragmentos abajo en la tabla y luego ofrecer una posible explicación del origen de estos errores.

Tabla 58: Muestras de los errores desde el aspecto sintáctico en nivel B2

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0260	nuevo delhi hay	famosa de las comidas	ejemplo chole bhature,	ord
ELEME0260	Tomb etcetera Hay	famoso de Mercado y supermercado	esta situado de la	Conc num
ELEME0143	profesores están muy	preocupado	con sus estudiantes.	Conc num
ELEME0237	hay muchos	centro comercial	. Y, por	Conc num
ELEME0144	gente y su	felicidades	dando la esperanza de	Conc num
ELEME0223	Ambas obras son	musical	o mejor digo que	Conc num
ELEME0260	Hay Muchas actividades	disponible	de la India sporte	Conc num
ELEME0260	etcetera . Hay	Mucahs famoso de Los monumentos	de la India TajMahal	Conc gen
ELEME0258	buenisima. Estoy	seguro	de que lo disfrutará	Conc gen
ELEME0262	.Yo tengo	la	Problema usar mucho pasar	Conc gen
ELEME0263	aplicaciones diferentes,	incluido	el servicio al cliente	Conc num Conc gen

Nota. Elaboración propia

En la muestra ELEME0260, se evidencia un orden sintáctico confuso con varios elementos adicionales y mal posicionamiento del adjetivo, por ejemplo, el uso de adjetivo antes del sustantivo, el uso del artículo después del adjetivo. Otro error común en el aspecto sintáctico es el de la concordancia de número. Se observa que hay una falta de concordancia entre adjetivos (con función de atributo) y sustantivos, como en el caso de ELEME0143. Algo similar sucede en la muestra ELEME0144, ‘la gente’ se considera como un sustantivo singular con un significado plural y se intenta concordar con un adjetivo plural: ‘felicidades’. Este error podría surgir desde la confusión entre la morfología y el significado de la palabra. Consiguientemente, desde las muestras disponibles en la tabla 58, no se puede ver una concordancia entre ‘los monumentos’ y ‘muchas’, ‘seguro’ con el sujeto que era femenino, el artículo ‘la’ con ‘problema’. Es de señalar, el error de ‘seguro’ se puede retomar como el error prototípico de concordancia entre los estudiantes de género femenino que tienden a usar la flexión masculina del adjetivo.

A continuación, se enfatiza en los errores de la morfología a nivel palabra en nivel B2. Después del análisis de datos se evidenció que todos los errores son de falsa selección de la morfología. Muchas veces estos errores tienen raíces en la lengua inglesa u otras lenguas habladas por el participante. En total, había 11 errores de falsa selección de la morfología. Abajo se exponen algunos ejemplos.

Tabla 59: *Muestras de los errores de la morfología en nivel B2*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0143	Hoy en día tenemos computanos, celulares,	tablet	y todo lo que podemos usar durante estudiar.	Fs morf
ELEME0143	antes no hicieron	casos	pero ahora todo el	Fs morf

ELEME0145	No tienen nada que comer y quedar muy	mucho	problema en (ilegible)	Fs morf
ELEME0238	Aquí todo sigue	igual	,el mismo rollo	Fs morf
ELEME0255	y necesito una	asistencia	por eso. Me	Fs morf

Nota. Elaboración propia

Se observa que muchas veces hay una confusión entre elementos que naturalmente son plurales por la locución o por la estructura morfosintáctica de la frase, por ejemplo: ‘caso’ en ‘hacerle caso’ es siempre singular. Asimismo, ‘mucho problema’ aunque existe gramaticalmente, no es adecuado usar desde el significado en el caso de la muestra ELEME0145 porque en este contexto se habla de varios problemas que surgen de no tener comida. Por eso, ‘muchos problemas’ es una combinación más frecuente y adecuado en su forma plural. En la muestra ELEME0238, se puede pensar que ‘igual’ podría ser un error de tipeo, pero se repitió el mismo error en la misma muestra y después de revisar el teclado en español e inglés, tanto del móvil como del portátil, no es posible por la distancia entre las dos letras. Al final, en el caso de ‘la asistencia’, se puede decir que es un uso inusual, quizás por los falsos cognados de la lengua inglesa. El estudiante quizás quería decir – ‘una ayuda’ o ‘un ayudante’ entre otras posibilidades.

7.3.2 Criterio etiológico

Desde el criterio etiológico, como se había observado en otros niveles, se analizaron los errores desde la interferencia de hindi como errores interlingüísticos y la hipercorrección, la simplificación y la sobregeneralización desde el criterio intralingüístico.



Figura 19: Errores desde el criterio etiológico en nivel B2

En primer lugar, hay 58 errores interlingüísticos (65,91%) y 30 errores intralingüísticos (34,09%). Dadas estas cifras, primero, se abarcan los errores interlingüísticos y luego los intralingüísticos para sacar algunas conclusiones y razonar los errores.

Tabla 60: Muestras de los errores desde la transferencia de hindi en nivel B2

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0040	actividades, transportes,		clima etc por eso	ETIO LM
ELEME0143	que podemos usar durante		estudiar. Personalmente,	ETIO LM
ELEME0208	enseñar a los niños		materias académicas. Las	ETIO LM
ELEME0210	entre los desarrolladores y		usuarios. En conclusión	ETIO LM
ELEME0210	la inteligencia artificial y		herramientas como ChatGPT,	ETIO LM
ELEME0223	. La cultura,		comida , festival,	ETIO LM

ELEME0228	, a menudo eclipsando		interacciones reales. Para	ETIO LM
ELEME0237	comida callejera junto con		restaurantes de alta gama	ETIO LM
ELEME0238	donde vende comida de		calle y es muy	ETIO LM
ELEME0238	punto tu sabes de		comida india mas que	ETIO LM
ELEME0259	con la ayuda en		casa, disfrutando momentos	ETIO LM
ELEME0260	nuevo delhi hay	famosa de las comidas	ejemplo chole bhature,	ETIO LM
ELEME0260	actividades disponible de	la India deporte,	cantar y bailar Y	ETIO LM
ELEME0260	y bailar Y	famoso de las lugares	de la India ejemplo	ETIO LM
ELEME0262	.Yo tengo	la	Problema usar mucho pasar	ETIO LM

Nota. Elaboración propia

Como se observa en la *tabla 60* y también en los resultados de los errores lingüísticos, la mayoría de los errores provienen de la omisión del artículo. Dado que, en hindi hay una ausencia completa de los artículos, se puede decir que hay una concordancia entre los resultados presentados y la dificultad lingüística prevista por parte de los participantes.

En el caso de la muestra ELEME0260, hay errores del orden sintáctico y la concordancia de número. Estos errores se pueden explicar desde la interferencia de hindi porque generalmente el adjetivo está puesto antes del sustantivo. Lo mismo sucede en el siguiente ejemplo, en el cual en vez de decir ‘deporte indio’, se dijo ‘Indian deporte’ porque en hindi se dice en el mismo orden sintáctico ‘भारतीय खेल/ *bhāratīya khēla*’. El mismo error se puede explicar desde la interferencia de inglés porque se dice ‘indian sports’ muchas veces sin traducción al hindi. En la siguiente oración: ‘famoso de las lugares’, hay un error de orden sintáctico que sigue la misma explicación de los ejemplos anteriores. Este error es de concordancia de número en la cual el artículo ‘las’ no concuerda con ‘lugares’. Puesto que en hindi ‘lugar’ es un sustantivo femenino y en español masculino, el estudiante extrapoló la regla de hindi a español, como ha pasado en otros niveles anteriores

bajo análisis. Algo similar sucede con el error ‘la problema’, además de ser una excepción en español (un sustantivo masculino que termina en –a), en hindi ‘problema’ tiene varios sinónimos con diferentes géneros como *samkaṭa* /संकट, *musibata* / मुसिबत, entre otras palabras. Entre estas dos palabras, ‘*samkaṭa*’ es masculino y ‘*musibata*’ es femenino. Dadas estas explicaciones, se puede decir que este error puede ser una confusión en la regla general de género en español o una interferencia de hindi desde la palabra ‘*musibata*’.

Tabla 61: *Muestras de los errores intralingüísticos en nivel B2*

Número de muestra	Texto antes	Error	Texto después	Código de error
ELEME0143	Es que la educación en el contexto de mi país,, no era la educación	más	importante como hoy	ETIO SIM
ELEME0143	podemos usar durante	estudiar	. Personalmente, cuando	ETIO SIM
ELEME0259	que es tranquilo.		Mamá manejar desayuno y	ETIO SIM
ELEME0259	algunas tareas y ayudar		mamá con cosas alrededor	ETIO SIM
ELEME0144	gente y su	felicidades	dando la esperanza de	ETIO SOBRE
ELEME0144	ellos no son	conciente	de las corcicuencias que	ETIO SOBRE
ELEME0145	. Necesitamos ser	abierto	para que adaptemos al	ETIO SOBRE
ELEME0207	Oro. Quedé	fascinado	con la riqueza histórica	ETIO SOBRE
ELEME0223	Ambas obras son	musical	o mejor digo que	ETIO SOBRE
ELEME0225	hotel son muy	amable	y siempre están disponibles	ETIO SOBRE
ELEME0228	La dopamina liberada por		likes y comentarios refuerza esta conexión, afectando	ETIO HIPER

			la concentración y la autoestima.	
--	--	--	-----------------------------------	--

Nota. Elaboración propia

Desde una perspectiva intralingüística, hubo 5 errores de simplificación, 24 errores de sobregeneralización y 1 error de hipercorrección. En el caso de la simplificación, los participantes usaron el adjetivo comparativo ‘más’ en lugar de ‘tan’ para crear una oración más sencilla. El participante debería haber omitido el sintagma directamente para dejar la oración final como “Es que la educación en el contexto de mi país, no era *tan* importante como hoy pero a mi me gustaría describir”. Aunque el estudiante eligió repetir el sustantivo para simplificar la oración. Otros ejemplos serían el uso del verbo ‘estudiar’ en lugar de ‘los estudios’ y la omisión completa de posesivo ‘mi’ antes de mamá en la muestra ELEM0259 (tabla 61). El error del uso del verbo ‘estudiar’ puede que tenga raíces en las preposiciones porque si se cambia ‘durante’ por ‘para’, es posible la combinación ‘para estudiar’. Por lo tanto, como la lengua no se puede aislar desde solo un sistema, este error se puede explicar desde dos paradigmas: nominal y preposicional. En cuanto al error de la omisión de ‘mi’ con mamá, se puede deducir la estrategia de evasión del uso de algún elemento por un posible miedo de cometer un error o una confusión entre varios elementos. En cuanto a la sobregeneralización, hubo errores de concordancia de número y género como en las muestras ELEM0144, ELEM0145 entre otras. Se puede ilustrar de estas muestras que los participantes usaron estas palabras erróneas desde su uso frecuente en la lengua en niveles básicos, por ejemplo, ‘felicidades’ para felicitar a alguien, ‘conciente’ de ‘ser consciente’, ‘abierto’ de ‘ser o estar abierto’ entre otros. Es de señalar que se observa que los participantes sobregeneralizan los componentes gramaticales de su flexión singular como es el caso de ‘abierto’. Así, no consideraron los sustantivos con los cuales tenían que concordar estos adjetivos. Por último, la omisión del artículo en la

muestra ELEM0288 se puede explicar desde la hipercorrección porque el participante corrigió innecesariamente un artículo antes de ‘likes’.

Transición de nivel A2 a B2

Después de observar los datos de cada nivel, se exponen las categorías y los errores que muestran el desarrollo de la interlengua de los estudiantes de nivel A2 a B2.

Como se evidenció previamente, desde el criterio lingüístico, la categoría gramatical fue la que más errores tuvo en todos los niveles. Abajo se muestra esta transición por niveles.



Figura 20: Transición de los errores lingüísticos de nivel A2 a B2

Se puede observar que los errores en la categoría gramatical suben de nivel A2 a nivel B2. En el caso de otras categorías, los errores en el aspecto sintáctico bajan de nivel A1 a B1 y luego suben otra vez en nivel B2. Asimismo, los errores de morfología siguen el mismo patrón con una bajada en nivel B1 y una subida a nivel B2. Esto se puede razonar también por las menores muestras disponibles en el caso de nivel B1. Aun así, con menores muestras en nivel B1, la categoría gramatical sigue un patrón ascendente y por eso se puede concluir que a los aprendientes de español en este nivel les dificultaron los temas gramaticales.

Con el fin de abordar los errores de la categoría gramatical, se expone la siguiente figura:

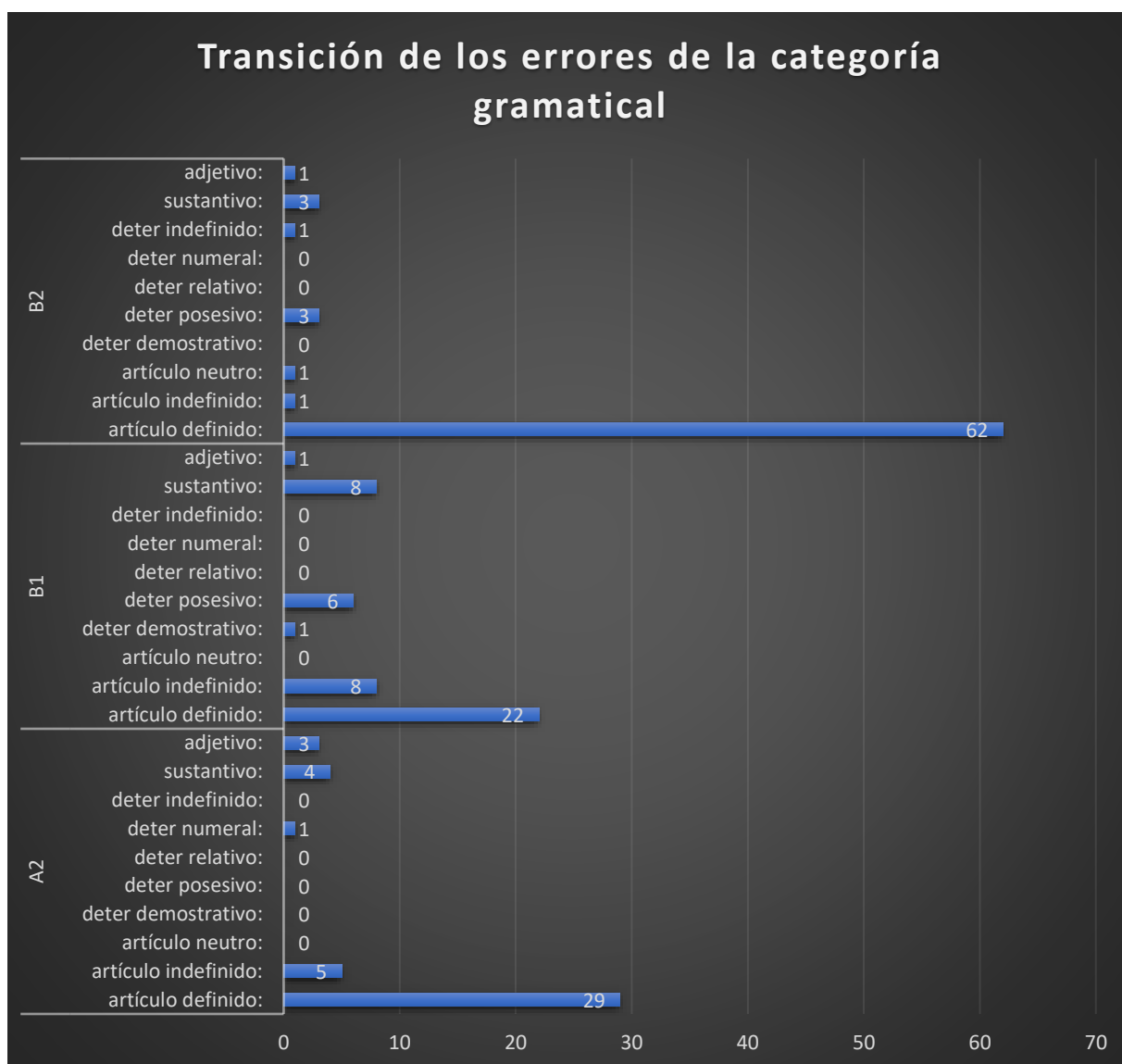


Figura 21: *Transición de los errores en la categoría gramatical desde nivel A2 a nivel B2*

En primer lugar, se puede observar que en nivel A2 solo hay 5 categorías que tienen errores, en nivel B2, hay 6 categorías con errores y en nivel B2, hay 7. Estos resultados de la ausencia de errores en ciertas categorías en ciertos niveles están de acuerdo con el concepto del desarrollo de la interlengua que explica cómo un estudiante durante el aprendizaje de un idioma crea varias hipótesis en niveles básicos, las aclara en niveles intermedios y las consolida y automatiza en niveles superiores. Es de destacar, que, en niveles básicos, los estudiantes no tienen muchas herramientas y/o recursos para explotar.

Sin embargo, en niveles intermedios y avanzados, los estudiantes pueden usar más componentes lingüísticos, en este caso, en el paradigma nominal.

En segundo lugar, el error de la omisión de artículo definido baja de nivel A2 a B1 y sube bastante en nivel B2. Como ya se ha explicado, esta bajada también se puede entender desde el número de muestras en nivel B1, pero también se puede explicar desde AE, AC y la IL. Estas teorías mencionan cómo los estudiantes al aprender un idioma crean hipótesis, las comprueban y las automatizan en sus interlenguas. Esta comprobación es evidente en nivel B1 porque hay muestra de alta frecuencia de errores en diferentes categorías de la taxonomía. Luego, los estudiantes ponen en práctica sus conocimientos para poder interiorizarlos en niveles más avanzados.

En tercer lugar, los estudiantes en nivel intermedio producen más texto o contenido, dado que tienen más recursos y herramientas. Por eso, el aumento de errores se refleja en nivel B1 en las categorías como: sustantivo, determinante posesivo y artículo indefinido que se aprenden a partir de niveles A1 y A2. Esto muestra la comprobación del funcionamiento de las reglas gramaticales que los estudiantes aprendieron en niveles básicos. Además, ellos suelen tener una dificultad con la elección entre los posesivos y los artículos por la ausencia completa de los artículos en hindi y un mayor uso de los posesivos. De todas formas, dado los datos y el aumento en los errores con el cambio de niveles, puede que los participantes no hayan tenido un proceso de corrección. Por lo tanto, hay una posibilidad de fosilización de los errores de niveles básicos en los niveles intermedios.

Ahora bien, desde el criterio etiológico, los errores mostraron la siguiente transición:

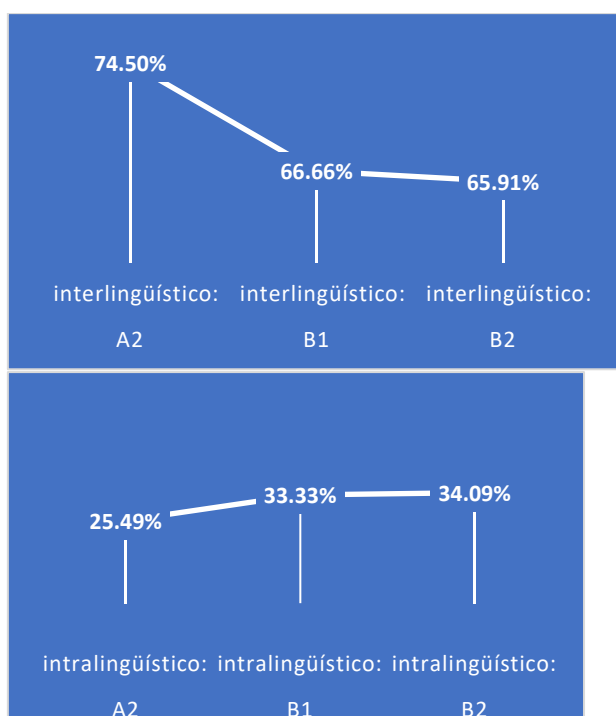


Figura 22: Transición de los errores interlingüísticos e intralingüísticos de nivel A2 a B2

Desde nivel A2 hasta nivel B2, los errores intralinguales suben y los errores interlinguales bajan. Esto quiere decir que la interferencia de español sube y la interferencia de hindi baja con el avance en el nivel. En esta figura, es de destacar que los errores intralinguales (25.50%) eran menos que los errores interlinguales (74.50%) en nivel A2. Estas cifras altas de errores interlinguales se mantuvieron hasta nivel B2 (34 % errores intralinguales y 66% errores interlinguales). En este orden de ideas, se puede decir que la interferencia de hindi, aunque baja en nivel B2, necesita una atención especial para evitar las fosilizaciones de errores.

Para resumir, según los resultados iniciales de esta investigación, en primer lugar, el paradigma nominal tuvo el máximo número de errores, dado que la ausencia de los artículos y la diferencia en la flexión de género en hindi y español se les dificultó a los estudiantes en todos los niveles. En segundo lugar, la omisión de los artículos muestra la ausencia de los elementos en hindi que existen en español. En tercer lugar, hubo varios errores desde la interferencia del inglés porque los estudiantes indios hablan inglés como segunda lengua y son competentes en este idioma. Se encontró que hubo 17 errores en

nivel A2, 4 errores en B1 y 10 errores en nivel B2. No obstante, el alcance y objetivo de este estudio no era analizar la interferencia del inglés y por eso no se incluyó esta categoría en la taxonomía adaptada ni se analizaron a profundidad estos errores.

8. Discusión

Para responder a la pregunta y el objetivo de la investigación y por ende hacer el análisis de errores morfosintácticos en el paradigma nominal en la población india con estudiantes hindi hablantes de nivel A2, B1 y B2, esta sección se dedica a explicar los resultados del análisis y compararlos con lo que encontraron otros investigadores.

Desde lo teórico, se encontraron dos referentes importantes, Díaz (2023) y Gadre (1996). En este estudio, se tomaron y actualizaron varias descripciones de los componentes gramaticales desde el paradigma nominal. El resultado en cuanto a los artículos y al error más común en esta categoría, en todos los niveles, es de la omisión. Esto, está en concordancia con estos autores que explicaron la ausencia de los artículos en hindi. Ellos mencionaron que los posesivos juegan un papel más importante en hindi que en español. Por eso, se puede decir que este componente lingüístico funciona, muchas veces, como el artículo en español.

También, se sustenta el conflicto entre el artículo definido e indefinido y cómo a los estudiantes indios con lengua materna hindi, les cuesta más el artículo definido que el artículo indefinido. Es porque a veces hay una correspondencia entre el número *ek* (1) de hindi y *un* (artículo indefinido) de español¹⁶. Por tanto, este elemento gramatical suele pasar como imperceptible o de baja importancia al crear una oración por la interferencia de hindi. No obstante, la realidad gramatical del español es diferente y los artículos son uno de los elementos más importantes en el paradigma nominal, dado que acompañan a los sustantivos en el sintagma nominal y al ser determinantes, le ayudan al interlocutor y al

¹⁶ Refiere a página 42

hablante a delimitar o identificar el sustantivo. Por ejemplo, introducir información nueva con los artículos indefinidos y referenciar algo conocido por los hablantes con los artículos definidos. Algo similar sucede con el estudio de López (2012) que encontró que la mayoría de los errores en el corpus de sus estudiantes serbio hablantes era el de omisión del artículo (36%), puesto que, no existe este componente gramatical en el serbio tampoco. En el presente estudio, se encuentra que la mayoría de los errores también son de este tipo.

Estos resultados se pueden extrapolar a la propuesta inicial de los conductistas Lado (1957) sobre la hipótesis de AC y Fries (1945) sobre la producción del material didáctico basado en un análisis contrastivo de la lengua materna del estudiante y la lengua meta. En este sentido, el análisis de la interlengua de los estudiantes y su desarrollo a lo largo de los diferentes niveles se puede explotar para entender sus necesidades. Asimismo, se promueve la creación del material didáctico y se recomienda tomar en cuenta errores comunes como la omisión del artículo, la interferencia de hindi en la concordancia de género y el orden sintáctico entre otros y su posible fosilización entre los estudiantes hindi hablantes. Esta perspectiva progresiva hacia el análisis contrastivo ayudaría tanto a los docentes trabajar con los estudiantes con el componente de miedo y reflexionar en los errores para poder autocorregirse.

Desde el criterio etiológico, en los resultados de este estudio se encuentra que en niveles A2, B1 y B2 más del 65% de errores tienen un origen interlingüístico. Si se estudia la transición de los errores de nivel A2 a B2, se puede evidenciar que los errores interlingüísticos son los máximos en nivel A2 (74,50%) y siguen bajando a nivel B1 (66,66%) y B2 (65,91%). En esta línea, se puede esperar una disminución de errores en niveles avanzados y así una menor interferencia de la lengua materna. Sin embargo, el porcentaje de los errores interlingüales sigue siendo más de 50% en nivel B2. Por esta razón, se puede decir que los resultados de esta investigación y de Paniagua (2005) con sus

estudiantes arabófonos de nivel avanzado (encontró que 3 de cada 5 errores tenían un origen intralingüístico), no están en concordancia y los aprendientes de español hindi hablantes tienen una alta interferencia de su lengua materna en sus interlenguas hasta nivel B2. Ahora bien, esta interferencia también podría ser errores que no han sido trabajados en niveles básicos y se van fosilizando en niveles intermedios y avanzados.

De modo idéntico, se encontró que los estudiantes solían relacionar el género de la palabra con el correspondiente al de su lengua materna. Cabe aclarar que, en hindi, aunque no haya artículos, los sustantivos se declinan según el género. Como consecuencia a este contraste lingüístico, a los hindi hablantes se les dificulta aprender los géneros de los sustantivos en español. Como una posible solución, los aprendientes usan el género de los sustantivos del hindi en español, lo que se evidencia en los resultados de esta investigación. De hecho, se puede ejemplificar este error en todos los niveles en este estudio con el sustantivo ‘lugar’ porque los participantes lo consideraron como un sustantivo femenino. Castenado (2019) en su estudio con ruso hablantes llegó a la misma conclusión porque sus estudiantes también flexionaron los géneros de los sustantivos en español según el género en ruso.

Como se ha visto en los resultados, los errores interlingüísticos bajaron y los errores intralingüísticos aumentaron hacia niveles avanzados. Sin embargo, los errores interlingüísticos tuvieron la frecuencia más alta, por tanto, la interferencia de hindi necesita un tratamiento especial por parte del docente. En relación con el punto anterior, la teoría de la interlengua (Selinker, 1972) y su característica de la variabilidad (Ortega, 2014) respalda el carácter transitorio del aprendizaje de una lengua, en este caso, de niveles básicos hacia niveles avanzados. Los resultados de Alexopoulou (2010) en su estudio con aprendientes de español griegos concuerdan con los resultados de esta investigación. En el

caso de presente estudio, se puede observar el mismo patrón de errores de niveles básicos hacia niveles intermedios.

Como una característica destacada de los participantes hindi hablantes del CAELE/2, ellos eran estudiantes activos de español o en proceso de inmersión cultural en países hispanohablantes como Colombia y España durante la participación en el estudio. Por lo tanto, se puede decir con certeza que estaban expuestos a la lengua española pasivamente porque estaban escuchándola diariamente en su ambiente. No obstante, no se puede decir si los estudiantes cometerán más errores de tipo interlinguales o intralinguales después de terminar sus estudios o al volver a sus países de origen. De acuerdo con los resultados de esta investigación, se espera que los errores interlingüísticos disminuyan al llegar a niveles más avanzados. Aunque, Liu (2012) en su estudio con aprendientes chinos, concluyó que los errores interlinguales aumentan después de terminar sus estudios académicos en China. Dado este antecedente, no se puede asegurar el cambio en el patrón de los errores interlinguales en la interlengua de los hindi hablantes después de terminar sus estudios o volver a sus países.

En otro orden de ideas, hubo varios errores de concordancia de género y número. En esta investigación, en los niveles A2, B1 y B2 hubo 16, 8, 22 errores de concordancia respectivamente. Además, los aprendientes en todos los niveles tienden a flexionar las palabras en masculino por defecto. Esto es aun más evidente entre las muestras escritas de las participantes de género femenino porque usan adjetivos con flexión masculina para ellas mismas. Estos errores son comunes en nivel A2 y se podrían fosilizar en niveles intermedios y avanzados (como se puede observar en muestras escritas de niveles B1 y B2). Dadas estas tendencias en los errores, se recomendaría a los docentes trabajar con los estudiantes a partir del contraste entre español e hindi y a la anticipación de errores.

Barrera (2023), en su estudio con noruegos, también concluyó que a los estudiantes se les dificultaba el concepto de la concordancia de género.

Con respecto a los argumentos presentados, todos los errores anteriores se pueden explicar desde la teoría de la interlengua (Selinker, 1972), en la cual se menciona cómo el estudiante crea varias hipótesis acerca de la lingüística de la lengua, las comprueba, las actualiza y las adquiere. En el mismo orden de ideas, se puede decir que el estudiante actualiza la interlengua en casi todos los niveles, sin embargo, en algunos niveles este proceso de actualización es más evidente. Por ejemplo, el estudiante aprende varios elementos nuevos de la lengua en niveles A1, A2, los comprueba y los actualiza en niveles A2 y B1 y si no hay una intervención o corrección guiada, las probabilidades de la fosilización de estos errores son altas en niveles B2 y C1.

En cuanto a esta investigación, este proceso del desarrollo de la interlengua se evidencia desde nivel A2 hasta nivel B2, especialmente en el caso de la interferencia de la lengua materna: hindi. En este sentido, en nivel A2, no hay muchos errores en la categoría de los determinantes porque en este nivel los estudiantes empiezan a aprender esta categoría gramatical y por eso, la ausencia o falta de errores está justificado. De la misma manera, hay más errores de adjetivos y sustantivos en este nivel por la confusión en el orden sintáctico y el mayor uso de estos elementos básicos para la construcción de una oración. En nivel B1, los errores, en general, bajan en frecuencia por la posible comprobación de las reglas y en nivel B2, se puede ver la fosilización de omisión de artículo (máximo número de errores) y errores en la concordancia de género. Se puede evidenciar estos errores como fosilizados porque los aprendientes ya llevan tres niveles aprendiendo este componente lingüístico. Además, durante estos niveles tienen varias oportunidades de comprobar y mejorar el error. Lo mismo se puede decir para otros

componentes que se aprenden en niveles básicos y siguen siendo una confusión en niveles avanzados.

Ahora bien, se puede decir que la interferencia de hindi es más alta en niveles A2 y B2 por dos razones diferentes. En nivel A2 los estudiantes suelen usar la lengua materna como un medio para aprender la lengua meta y la interferencia es consciente e intencional para crear un nuevo conocimiento con una alta tendencia de autocorrección. No obstante, en nivel B2, la interferencia es inconsciente y no intencional porque los estudiantes tienen la tendencia de fosilizar los errores y no reflexionar sobre ellos por la falta de la capacidad de identificar el error (Rodrigo, 2018). Desde una perspectiva pedagógica, esta tendencia de errores podría estar dada por la falta de un abordaje adecuado en el aula, un trabajo significativo que ayudaría a los estudiantes entender tanto el significado como la forma de los componentes gramaticales en el paradigma nominal.

Adicionalmente, en cada nivel hubo casos de muestras perfectas o con cero errores, específicamente, en A2, hubo 7 muestras escritas, 2 casos en B1 y 6 en B2. La mayoría de estas muestras son de 1-2 líneas o tiene errores en otros paradigmas como preposicional o verbal. Por lo tanto, se puede decir que estaban escritas por humanos. Sin embargo, hubo muestras como ELEME0219 y ELEME0227, las cuales se presume que están escritas por la inteligencia artificial (IA) porque no se encontraron errores ortográficos, pragmáticos, ni discursivos y hubo una falta de personalización. Además, al comparar esta muestra con otros textos que escribió el mismo estudiante, se puede diferenciar entre la muestra escrita por él/ella y la muestra escrita con la ayuda de IA. En otros casos, como en ELEME0209 y ELEME0229, es incierto el uso de IA porque hay opiniones y comentarios personales.

Para concluir, los resultados de esta investigación también están de acuerdo con Chomsky (1965) y su teoría de cognitivismo que considera el error como un proceso

natural y una muestra de progreso y crecimiento porque el desarrollo de la interlengua muestra un progreso y actualización de ella. De la misma manera, esta investigación hace una reflexión desde AC, AE e IL con una perspectiva progresiva y positiva (ver el error desde un síntoma de progreso y no de fracaso) para poder etiquetar a los errores de los estudiantes hindi hablantes y analizar tanto sus errores como su interlengua. Esto ayudaría a los aprendientes actualizar la interlengua y evitar fosilizaciones en niveles avanzados.

9. Conclusiones

Al inicio de este documento se propone que el objetivo general de esta investigación es analizar los errores morfosintácticos en el paradigma nominal de estudiantes de español hindi hablantes en niveles A2, B1 y B2, en muestras escritas del corpus CAELE/2, desde la perspectiva de AC y AE.

Mediante al análisis de estas muestras escritas y los hallazgos del estudio, se propone una serie de conclusiones.

- La primera conclusión a la que se llega es que lo que más se le dificulta a nivel morfosintáctico a los aprendientes de español hindi hablantes es el paradigma nominal. Esto se evidencia en que la mayoría de los errores se encontraron en este paradigma y el error más común fue la omisión del artículo que atravesó a todos los niveles. Esto se puede explicar desde las diferencias lingüísticas entre español e hindi, aun así, las dos lenguas tienen muchas similitudes como la flexión de género y número, por lo tanto, se recomienda a los profesores didactizar las clases desde una explotación lingüística de los conocimientos previos y luego agregar nuevos componentes.
- En cuanto a la literatura encontrada, no hay muchos estudios en la enseñanza y la didáctica de la lengua española en India. Con el fin de promover más investigaciones en este campo de estudio y proponer material didáctico para la

población india, se pueden consultar los resultados de esta investigación y el corpus CAELE/2.

- En cuanto a los aspectos metodológicos, en primer lugar, el cuestionario sirvió con eficacia en la recopilación de corpus en formato virtual por *Google forms*. En segundo lugar, la taxonomía adaptada de Ferreira y Elejalde (2020), modificada después de la validación por tres calificadoras, logró etiquetar todos los errores desde el criterio lingüístico y etiológico. En tercer lugar, el software *UAM corpus tool* ayudó y apoyó la anotación de las muestras escritas y el respectivo análisis de datos con éxito. Dicho esto, este procedimiento comprobado con el juicio de expertos puede favorecer a las investigaciones de la misma línea para llegar a resultados significativos.
- En cuanto a los resultados, uno de los errores más frecuentes fue el error de la concordancia de género. Los participantes (sea mujer o hombre) usaron la flexión masculina tanto con los adjetivos que representaron a ellos mismos como con los sustantivos femeninos. Dado estos errores, se recomienda a los docentes, primero, estar atentos a los errores de la concordancia de género desde la flexión masculina y corregirlos en el momento adecuado. Segundo, enseñar los sustantivos con diferentes combinaciones de determinantes y no solo como léxico aislado. Por ejemplo, combinar los sustantivos con los posesivos, los artículos etc.
- Los resultados desde el criterio etiológico mostraron una clara interferencia de hindi en el aprendizaje de español en niveles A2, B1 y B2. Por eso, se recomienda que los docentes promuevan espacios de reflexión para hacerles conscientes a los estudiantes sobre el paradigma nominal de su lengua materna hindi a través de preguntas orientadoras y usarla como una herramienta para facilitar el aprendizaje de español. Es de aclarar que esta investigación no promueve la traducción de las

lenguas sino, la reflexión personal del paradigma nominal de hindi por parte de los estudiantes.

- A partir de los resultados y la discusión se sugiere que el docente sepa los rasgos lingüísticos básicos del hindi para poder entender las necesidades de los estudiantes hindi hablantes y facilitar el proceso de aprendizaje de español.
- El análisis de errores desde las teorías de análisis contrastivo (interlingüístico), análisis de errores (intralingüístico), la interlengua y la metodología de la lingüística de corpus brindó un análisis holístico y completo. Estas perspectivas ofrecieron resultados desde diferentes miradas y entender mejor la raíz del error.

Propuestas para futuras investigaciones

- Esta investigación alimentó al corpus CALELE/2 para la población india. Este corpus hasta ahora ha recopilado 81 muestras de diferentes lenguas maternas originarias de India y cada año se agregarán más perfiles para ampliar el corpus y ofrecer más herramientas a los profesores de español en India. Se puede explorar y explotar este corpus para hacer un análisis de necesidades de los estudiantes hindi hablantes y/o didactizar y llevarlo al aula para hacer correcciones guiadas como una actividad.
- En esta investigación se hizo un análisis profundo del paradigma nominal a nivel morfosintáctico de la lengua española, aunque, se encontró que hubo muchos errores en el paradigma preposicional y verbal. Este análisis también requerirá un estudio contrastivo entre hindi y español. A partir de estos vacíos, surge la necesidad de llevar acabo estos estudios contrastivos.

- Durante la anotación de los errores, se encontraron varios errores que tenían interferencia del inglés. Se puede explorar la interlengua de los estudiantes desde la interferencia de inglés y la lengua materna de los aprendientes de español en India.
- Con base en esta investigación, se promueve la creación y la adaptación de material didáctico para los estudiantes indios, en este caso hindi hablantes. Es de destacar que el material existente para esta población es escaso o desactualizado.
- Desde la metodología, se puede retomar la lingüística de corpus y la taxonomía adaptada para etiquetar los errores. Se sugiere adaptar la taxonomía y la metodología mixta para hacer análisis de errores en el futuro e investigar otras dimensiones de la lengua.
- En cuanto las recomendaciones hechas, los mismos resultados y el análisis contrastivo hecho entre hindi y español se pueden extrapolar a la lengua urdu porque las dos lenguas tienen la misma gramática.

Limitaciones

- Durante la validación del instrumento, no se pudo encontrar calificadores con conocimiento lingüístico de la lengua hindi. Por esta razón, solo fue posible validar la taxonomía desde el criterio lingüístico.
- En cuanto a la literatura, no hay muchas investigaciones de análisis de errores en la India para poder contrastar los resultados con otros investigadores u otras lenguas. Tampoco existen referentes bibliográficos o investigaciones en cuanto al análisis contrastivo entre hindi y español. En este estudio, solo se pudo consultar la producción bibliográfica de Díaz (2023) y Gadre (1996). Otra fuente que se encontró en la bibliografía de estos libros fue los enlaces de Youtube que se pueden utilizar como una referencia.

- El análisis de errores presentó varios desafíos para el investigador porque el español no era su lengua materna. El mayor desafío era no marcar errores innecesarios tanto de forma como de fondo. La ayuda de sus pares y la asesora aclaró varias de estas dudas en la morfosintaxis del paradigma nominal.
- Este estudio no tomó en cuenta la interferencia del inglés por el alcance de la investigación.
- La IA era un factor emergente para analizar que no se esperaba, aunque se esperaba el posible uso de recursos y herramientas disponibles en línea. En este sentido, al final del cuestionario los estudiantes mencionaron que usaron varias herramientas como *Wordreference*, *Spanishdict* entre otras. Estas herramientas no interfieren en el análisis morfosintáctico porque la consulta era más de tipo léxico.

10. Referencias

- Alexopoulou, A. (2006). *Los criterios descriptivo y etiológico en la clasificación de los errores del hablante no nativo: una nueva perspectiva*. 17–35.
<https://digibug.ugr.es/handle/10481/30431>
- Asociación, A. A.-A. del X. C. de la, & 2010, undefined. (n.d.). Errores intralinguales e interlinguales en la interlengua escrita de aprendientes griegos de e/le.
Cvc.Cervantes.Es A Alexopoulou Actas Del XVI Congreso de La Asociación Internacional de Hispanistas, 2010•cvc.Cervantes.Es. Retrieved February 6, 2024, from https://cvc.cervantes.es/literatura/aih/pdf/16/aih_16_2_004.pdf
- Azevedo, M. M. (1980). The interlanguage of advanced learners: an error analysis of graduate students' Spanish. *IRAL: International Review of Applied Linguistics in Language Teaching*, 18(3), 217.
- Baralo, (2004) La interlengua de hablante
https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=Baralo%2C+2004+la+interlengua+de+hablante&btnG=
- Bernal Chávez, J. A, & Hincapié Moreno, D. A. (2018). *Lingüística de corpus*. Instituto Caro y Cuervo. Imprenta Patriótica.
- Bharti, P. (2019). La enseñanza de español como lengua extranjera en la India. *Matrices En Lenguas Extranjeras*, 12. <https://doi.org/10.15446/male.n12.76970>
- Bru, A., R., M., & 2016, undefined. (2016). Calcos sintácticos y errores ELE en el paradigma verbal en sujetos bilingües en Barcelona. *Revistas.Nebrija.Com MRA Bru Revista Nebrija de Lingüística Aplicada a La Enseñanza de Lenguas, 2016*•revistas.Nebrija.Com. <http://revistas.nebrija.com/revista-linguistica/article/view/274>
- Cabrera, A. F., & Gómez, J. E. (2017). Análisis de errores recurrentes en el Corpus de Aprendices de Español como Lengua Extranjera, CAELE. *Revista Brasileira de Lingüística Aplicada*, 17(3), 509–538. <https://doi.org/10.1590/1984-6398201710927>
- CLASE ELE Propuesta práctica, L. DE, & La Actividad, F. DE. (n.d.). *La comunicación no verbal en contextos y libros de ELE en India*. (n.d.).
- Corder, S. P. (1975). Error analysis, interlanguage and second language acquisition. Cambridge.Org. <https://www.cambridge.org/core/journals/language-teaching/article/erroranalysis-interlanguage-and-second-languageacquisition/5C9DFA83D40DBD644A6567B71FA9AD98>
- Corpus Linguistics: Method, Theory and Practice* - Tony McEnery, Andrew Hardie - Google Libros. (n.d.). Retrieved October 28, 2023, from https://books.google.com.co/books/about/Corpus_Linguistics.html?id=3j3Wn_ZT1qwC&redir_esc=y
- Correal, T. (2007). “Conducta verbal” de BF Skinner, 1957-2007.
<https://dialnet.unirioja.es/servlet/articulo?codigo=2540789>
- D. Wilkins, 1969 analysis
https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=D.+Wilkins%2C+1969+analysis&oq=D.+Wilkins%2C+1969+ana
- De Saussure, F. (1971). *Curso de lingüística general* (A. Alonso, Trad.). Editorial Losada, SA (Trabajo original publicado en 1916).
- Díaz, G (2023) Gramática de hindi y urdu para hispanohablantes.

- Diogenes, N. C. (1965). Persistent topics in linguistic theory.
<https://journals.sagepub.com/doi/abs/10.1177/039219216501305102>
- Durao, A.B 2007 *La interlengua*
https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=Durao%2C+A.B+2007+La+interlengua&btnG=
- Dušková, L. (1969). *On sources of errors in foreign language learning*.
<https://www.degruyter.com/document/doi/10.1515/iral.1969.7.1.11/html>
- E Tognini-Bonelli. (2001). Corpus linguistics at work.
<https://www.torrossa.com/gs/resourceProxy?an=5002198&publisher=FZ4850>
- Ellis, N. (2015). *Implicit AND explicit language learning Their dynamic interface and complexity*. <https://doi.org/10.1075/sibil.48.01ell>
- Ellis, R. (1994). *The study of second language acquisition*.
[https://books.google.es/books?hl=es&lr=&id=3KglibyrZ5sC&oi=fnd&pg=PP7&dq=%22The+Study+of+Second+Language+Acquisition%22+\(1994&ots=wGTpbm3yvU&sig=gJyVnEHSZHMf7_HYGPrZXCOYYX4](https://books.google.es/books?hl=es&lr=&id=3KglibyrZ5sC&oi=fnd&pg=PP7&dq=%22The+Study+of+Second+Language+Acquisition%22+(1994&ots=wGTpbm3yvU&sig=gJyVnEHSZHMf7_HYGPrZXCOYYX4)
- Fries, C. (1945). *Teaching and learning English as a foreign language*.
<https://eric.ed.gov/?id=ED071477>
- Gadre, V. G. (1996) Estructuras gramaticales de hindi y español.
- García Álvarez, A. (2015). *¿Y tú por qué estudias español? Una propuesta didáctica para el contexto multicultural y multilingüe en el sistema educativo superior de la India*.
- Gargallo, S. (1993). Análisis contrastivo, análisis de errores e interlengua en el marco de la lingüística contrastiva. *Cir.Nii.Ac.Jp*.
<https://cir.nii.ac.jp/crid/1130282273014023808>
- Gass, S., Behney, J., & Plonsky, L. (2020). *Second language acquisition: An introductory course*. <https://www.taylorfrancis.com/books/mono/10.4324/9781315181752/second-language-acquisition-susan-gass-jennifer-behney-luke-plonsky>
- Gomez, H. G. (n.d.). La teoría de análisis de errores y su aplicación en la enseñanza de las lenguas. *Revista Universidad Eafit*, 39–92.
- Hammarberg, B. (1974). The insufficiency of error analysis. *IRAL - International Review of Applied Linguistics in Language Teaching*, 12(1–4), 185–192.
<https://doi.org/10.1515/IRAL.1974.12.1-4.185/HTML>
- Haugen, E. (1953). The Norwegian Language in America, a Study in Bilingual Behavior, Volume 2. *The Norwegian Language in America, a Study in Bilingual Behavior, Volume 2*. <https://doi.org/10.9783/9781512820522/HTML>
- Helsinki declaration 1964*
https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=helsinki+declaration+1964+&btnG=
- Hernández, R. (2018). Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta Las rutas Cuantitativa Cualitativa y Mixta. *McGRAW-HILL Interamericana Editores S.A. de C.V.*, 753.
<http://repositorio.uasb.edu.bo:8080/bitstream/54000/1292/1/Hernández- Metodología de la investigación.pdf>
- Herrero, J. R. (n.d.). *Enseñar ELE en India*.
https://www.exteriores.gob.es/Documents/FichasPais/INDIA_FICHA%20PAIS.pdf
- Isabel, I., & Ochoa, E. (n.d.). *UNIVERSIDAD AUTÓNOMA DE BARCELONA FACULTAD DE FILOSOFÍA Y LETRAS DEPARTAMENTO DE FILOLOGÍA Análisis de errores de la interlengua fónica de coreanos aprendices de español*.

- J Norrish. (1983). Language learners and their errors. *Cir.Nii.Ac.Jp*.
<https://cir.nii.ac.jp/crid/1130282269212889728>
- Jaramillo, A. L. P. (2022). *Análisis comparativo de errores morfosintácticos del paradigma nominal entre muestras escritas y orales de aprendientes de español como lengua extranjera de nivel B2- Corpus CAELE/2*.
- Kalocsányiová, E. (2012). *El uso de cortometrajes en el aula de ELE: Una propuesta didáctica para el contexto multicultural de la India*.
- Kartik, V. (2021). La gamificación en el aula de ELE. *Universidad de Valladolid*.
- Kumar, M. (2018). Peculiaridades de enseñanza del español en la India y el uso del hindi en las clases. *IX Congreso Internacional de La Asociación Asiática de Hispanistas, Bangkok 2016*.
- Lado, R. (1957). *Linguistics across cultures; applied linguistics for language teachers*.
<http://thuvienso.thanglong.edu.vn/handle/TLU/6837>
- León Wintaco, L. D. (2019). *Análisis de errores morfosintácticos en un corpus escrito de aprendientes sordos*.
- Linguistics, J. F.-T. I. in C., & 1981, undefined. (n.d.). Theoretical issues in contrastive linguistics.
<https://www.torrossa.com/gs/resourceProxy?an=5063599&publisher=FZ4850>
- Liu, N. (2012). *Análisis de errores y evolución de la interlengua de los estudiantes sinohablantes: esbozo de un estudio longitudinal*. <https://ddd.uab.cat/record/106258>
- López, M. J. N., & 2012. (2012). Análisis de errores de interlengua de español para serbiohablantes: Dificultades de uso del artículo. *Educacionyfp.Gob.Es*.
<https://www.educacionyfp.gob.es/dam/jcr:6cb1f335-7421-46c6-8cb7-344a9e0489a9/2014bv1514mariajosenadal-pdf.pdf>
- Mackey, A., Al-Khalil, M., Atanassova, G., Hama, M., Logan-Terry, A., & Nakatsukasa, K. (2007). Teachers' Intentions and Learners' Perceptions about Corrective Feedback in the L2 Classroom. *International Journal of Innovation in Language Learning and Teaching*, 1(1), 129–152. <https://doi.org/10.2167/ILLT047.0>
- Mackey, A., Polio, C., & McDonough, K. (2004). The relationship between experience, education and teachers' use of incidental focus-on-form techniques.
<http://Dx.Doi.Org/10.1191/1362168804lr147oa>, 8(3), 301–327.
<https://doi.org/10.1191/1362168804LR147OA>
- Mancera, A. M., Martínez, I. P., Canales, A. B., Fernández, L. C., & Granda, J. S. (n.d.). Corpus para el análisis de errores de aprendices de E/LE (Corane). *Universidad de Alcalá*, 527–534.
- Manuel Criado de Val, *Fisonomía del idioma español*.
https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=Manuel+Criado+de+Val%2C+Fisonom%C3%ADa+del+idioma+espa%C3%B1ol%2C+Madrid.+Aguilar%2C+1957%2C&btnG=
- Martcorena Quintanilla, M., Alcina Franch, J., & Manuel Blecua, J. (2011). La gramática desde 1973 hasta 2010. *Ojs.Ucp.Edu.Pe*, 1(2), 142–148.
<https://doi.org/10.22386/ca.v1i2.17>
- Martín, J. L. O., & Djandué, B. D. (2014). El uso de la lengua meta en el curso inicial de Español Lengua Extranjera (E/LE) en Costa de Marfil. *Porta Linguarum*, 21, 153–164. <https://doi.org/10.30827/DIGIBUG.30488>
- Méndiz Noguerol, A. (2015). Estrategias de «city placement» (emplazamiento de ciudades en el cine) en la promoción del turismo español. El caso de Zindagi Na Milegi Dobara

- (Sólo se vive una vez, 2011). *Pensar La Publicidad. Revista Internacional de Investigaciones Publicitarias*, 8(2).
https://doi.org/10.5209/rev_pepu.2014.v8.n2.50734
- Muñoz-Basols, J., Moreno, N., Inma, T., & Lacorte, M. (2016). *Introducción a la lingüística hispánica actual: teoría y práctica*.
https://books.google.es/books?hl=es&lr=&id=Cn6_DQAAQBAJ&oi=fnd&pg=PT8&dq=Mu%C3%B1oz-Basols+introducci%C3%B3n+a+la+linguística&ots=Rz6DeXWVOD&sig=SFIQmYi7QU7tBB7hG-rNSiNBe_c
- Ortega, L. (2014). *Understanding second language acquisition*.
[https://books.google.es/books?hl=es&lr=&id=30bKAgAAQBAJ&oi=fnd&pg=PP1&dq=:+%22Understanding+Second+Language+Acquisition%22+\(2009\)&ots=Cd0_YXjHVq&sig=Q7yBN-6Fq7NMtZFib28cpKULir8](https://books.google.es/books?hl=es&lr=&id=30bKAgAAQBAJ&oi=fnd&pg=PP1&dq=:+%22Understanding+Second+Language+Acquisition%22+(2009)&ots=Cd0_YXjHVq&sig=Q7yBN-6Fq7NMtZFib28cpKULir8)
- Parodi, G. (2008). Lingüística de corpus: una introducción al ámbito. *RLA. Revista de Lingüística Teórica y Aplicada*, 46(1), 93–119. <https://doi.org/10.4067/S0718-48832008000100006>
- Piaget, J., Chomsky, N., & M Piattelli-Palmarini. (1979). Théories du langage, théories de l'apprentissage: le débat entre Jean Piaget et Noam Chomsky.
<https://cir.nii.ac.jp/crid/1130282268931953664>
- Pujol, Ó. (2012). *El español en el mundo*.
- Ramallo, F. C. (n.d.). *Reflexión en la enseñanza: creencias de seis profesores nativos de ELE en el contexto de India*.
- Rehman, A. (2021). La importancia de la enseñanza de la competencia pragmática en el curso del español con fines específicos para los guías turísticos en la India. *Revista Lengua y Cultura*, 2(4), 98–107.
<https://repository.uaeh.edu.mx/revistas/index.php/lc/issue/archive>
- Richards, J. C. (1974). Problems and prospects. *RELC Journal*, 5(2), 69–84.
<https://doi.org/10.1177/003368827400500207>
- Rodrigo-Mateu, A. (2018). La expresión escrita contextualizada y la corrección guiada en ELE. 1–30. http://chutneyconsalsa.blogspot.com.in/p/cuentos_1.html.
- Roger, L., Paniagua, R., & De Salamanca, U. (2001). Análisis de errores e interlengua de aprendices árabes de español como lengua extranjera. *Redined.Educacion.Gob.Es*.
<https://redined.educacion.gob.es/xmlui/handle/11162/76246>
- SAITO, A. (2002). *Análisis de errores en la expresión escrita de los estudiantes japoneses. Curso de traducción inversa japonés-español*.
- Sánchez Jiménez Dirigida por la Dra María del Carmen Fernández Juncal Salamanca, D. (2006). *Universidad de salamanca aprendizaje del español como lengua extranjera*.
- Sánchez-Jiménez, D. (n.d.). *El análisis de errores ortográficos de estudiantes filipinos en el aprendizaje de español como LE y su aplicación didáctica*.
<https://academicworks.cuny.edu>
- Scriba, A. S. (2018). *Préstamos remotos del sánscrito al español: el comercio de piedras preciosas en la india antigua*. 3.
- Selinker, L. (1972a). Interlanguage. *IRAL - International Review of Applied Linguistics in Language Teaching*, 10(1–4), 209–232. <https://doi.org/10.1515/IRAL.1972.10-1-4.209/MACHINEREADABLECITATION/RIS>

- Selinker, L. (1972b). Interlanguage. *IRAL - International Review of Applied Linguistics in Language Teaching*, 10(1–4), 209–232. <https://doi.org/10.1515/IRAL.1972.10.1-4.209/MACHINEREADABLECITATION/RIS>
- Silva, F. (2023). *Análisis de errores en la concordancia nominal del español. Un estudio sobre un grupo de estudiantes noruegos de ELE*. <https://munin.uit.no/handle/10037/30437>
- Singh, V. K. (n.d.-a). *El uso del texto literario en la clase de ELE: de la India a las indias*. <http://www.youtube.com/watch?v=pYq6B-ibjo><http://www.youtube.com/watch?v=ca5rBK4q7Kg><http://www.youtube.com/watch?v=QjtdbCj99QI><http://www.youtube.com/watch?v=iNlc1n3G7T85>. *Dinámica: Mixta*.
- Singh, V. K. (n.d.-b). *Los currículos de ELE en las universidades indias: falta de capacitación y formación docente en el área de ELE*.
- Singh, V. K. (n.d.-c). *¿Por qué el soporte audiovisual en el aula india de ELE?*
- Singh, V. K., & Ira, M. (2010). *El cine como instrumento didáctico en las aulas de ELE en un país de Bollywood*. 1–18. <http://redalyc.uaemex.mx/src/inicio/ArtPdfRed.jsp?iCve=92117173010>
- Sirinit, T. (n.d.). *Texto con elementos culturales (español-tailandés) / analysis of translation errors of thai students in the translation of a text with cultural elements (spanish-thai)*.
- Tobías, C. I. (2017). *La enseñanza del español a hablantes de hindi: de la competencia reflexiva a la actuación*.
- Torijano, J. A. (2014a). El análisis de errores de estudiantes lusófonos sobre los verbos españoles (I): presente y futuro. *Revista.Sel.Edu.Es*, 44(1), 145–178. <http://revista.sel.edu.es/index.php/revista/article/view/121>
- Torijano, J. A. (2014b). El análisis de errores de estudiantes lusófonos sobre los verbos españoles (i): presente y futuro. *rse*, 44(1), 145–178.
- Transliteration of [Devanagari](#) and related [Indic scripts](#) into [Latin characters](#)
<https://www.ashtangayoga.info/philosophy/sanskrit-and-devanagari/transliteration-tool/>
- Vega, C. C. (2019). *Análisis de errores durante la adquisición del español en rusohablantes de nivel B1 en el aula de ELE: el uso del artículo*. <https://repositorio.unican.es/xmlui/handle/10902/18002>
- Winke, P., Gass, S., & Sydorenko, T. (2013). Factors Influencing the Use of Captions by Foreign Language Learners: An Eye-Tracking Study. *The Modern Language Journal*, 97(1), 254–275. <https://doi.org/10.1111/J.1540-4781.2013.01432.X>
- Wintaco, L. L. (2019). *Análisis de errores morfosintácticos en un corpus de aprendientes sordos*. <http://bibliotecadigital.caroycuervo.gov.co/1314/>

11. Anexos

Anexo 1

Excepción en la flexión en masculino singular ante un sustantivo

Singular			Plural	
masculino pospuesto	masculino antepuesto	femenino	masculino	femenino
alguno	algún	alguna	algunos	algunas
bueno	buen	buen	buenos	buenas
malo	mal	mala	malos	malas
ninguno	ningún	ninguna	ningunos	ningunas
primero	primer	primera	primeros	primeras
tercero	tercer	tercera	terceros	terceras
uno	un	una	unos	unas

Nota. Adaptado de libro Estructuras gramaticales de hindi y español, Gadre (1996)

Anexo 2

El orden sintáctico en hindi y español: Comparación

Español	Hindi
---------	-------

Frase 1: Yo soy abogado. Frase 2: Ellas quieren agua. El orden sintáctico: Sujeto(S)+ Verbo(V)+ Objeto(O) No es necesario usar el sujeto en la oración ya que en el verbo queda explícito el uso de sujeto.	Frase 1: Mai vakil hun. Frase 2: Unhe pani chahiye. El orden sintáctico: Sujeto(S)+ Objeto(O)+ Verbo(V) Es importante y obligatorio usar el sujeto en la oración porque no se queda explícito en el verbo.
Frase3: Ellos corren. Frase 4: Yo como. El orden sintáctico: S+ V. Si el verbo es intransitivo, o no tiene complementos, el orden sintáctico es lo mismo en las dos lenguas.	Frase 3: We bhagte hain. Frase 4: Mai khata hun. El orden sintáctico: S+ V. Si el verbo es intransitivo, o no tiene complementos, el orden sintáctico es lo mismo en las dos lenguas.
Frase 7: Quiero a Khushbu. Frase 8: Nosotros enseñaremos a los niños. Orden sintáctico: S + V+ a(Preposición)+ objeto directo (OD) Frase 9: Vasant tomó el té. Frase 10: Los chicos condujeron el bus. Orden sintáctico: S + V+ OD En una oración transitiva con objeto directo, el OD va al final.	Frase 7: Mainn khushbu ko chahta huun. Frase 8: Ham bacchon ko sikhaaenge. Orden sintáctico: S + objeto directo (O.D) + ko (ERG)¹⁶+ V. Frase 9: Vasant ne chai pii. Frase 10: Ladkon ne bas chalai. Orden sintáctico: S + ne(ERG)+ OD+ V En una oración transitiva con objeto directo, el objeto directo se posiciona antes del verbo, aunque el posicionamiento del ergativo depende de su tipo.

Frase 11: Es un gran edificio. Frase 12: Es un buen hombre. Orden sintáctico: Verbo+ artículo+ adjetivo+ sustantivo. Frase 13: Es un edificio grande. Frase 14: Es un hombre bueno. Orden sintáctico: Verbo+ artículo+ sustantivo+ adjetivo. En español la posición de los adjetivos cambia el significado. Esta regla funciona con pocos adjetivos. Frase 15: Rohan es un hombre alto. Orden sintáctico: S+ verbo+ artículo+ sustantivo+ adjetivo. El adjetivo va después de sustantivo, al menos que son formas apocopadas. Frase 16: Este es un libro. Frase 17: Aquello es un carro. Orden sintáctico: Demostrativo+ verbo+ artículo+ sustantivo. Los demostrativos tienen por lo general, van primero.	Frase 11: Yeh ek badi imarat hai. Frase 12: Yeh ek acha aadmi hai. Orden sintáctico: Demostrativo+ ek(uno)+ adjetivo+ sustantivo+ verbo. Frase 13: Yeh imarat badi hai. Frase 14: Yeh aadmi acha hai. Orden sintáctico: Demostrativo+ sustantivo+ adjetivo+ verbo. En hindi, el significado de los adjetivos se mantiene, sin importar la posición (antes o después de los sustantivos). Frase 15: Rohan ek lamba aadmii hai. Orden sintáctico: S+ ek(uno) + adjetivo+ sustantivo+ verbo. En general, los adjetivos van antes de los sustantivos. Frase 16: Yeh ek kitaab hai. Frase 17: Weh ek gadi hai. Orden sintáctico: Demostrativo+ ek (uno)+ sustantivo+ verbo. Los demostrativos en hindi, por lo general, van primero.
Dificultades para estudiantes: Practicar el no uso de pronombres (Yo, él, ella, etc.) en español. En español el verbo concuerda con el Sujeto en número y persona, pero no en el género como en cado de hindi. El ergativo juega un papel importante en hindi y así nacen las declinaciones. En el orden sintáctico se nota que a veces, el ergativo es equivalente a la preposición en español, pero no es el caso siempre. Esta diferencia es de destacar en la clase con hindi hablantes. El posicionamiento del adjetivo a veces resulta difícil en español para los hindi hablantes por la misma distinción mostrada en las frases 15.	

Nota. Ilustración propia a partir de Gadre (1996) y Díaz (2023)

Anexo 3

Formulario para recopilar el corpus

<https://forms.gle/TauCXrjnn7oaShN29>



Anexo 4

Validación del instrumento

https://docs.google.com/document/d/17km9ADH04U51W-k-4j_33WZ91069kiHgkagXInMYqTY/edit?usp=sharing

