



# Los errores léxicos en textos escritos en español por alumnos universitarios brasileños en formación como profesores de Español Lengua Extranjera

Cynthia Baerlocher Rocha

**ADVERTIMENT.** La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX ([www.tdx.cat](http://www.tdx.cat)) i a través del Dipòsit Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA.** La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR ([www.tdx.cat](http://www.tdx.cat)) y a través del Repositorio Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING.** On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX ([www.tdx.cat](http://www.tdx.cat)) service and by the UB Digital Repository ([diposit.ub.edu](http://diposit.ub.edu)) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Tesi doctoral presentada per En/Na

**Cynthia BAERLOCHER ROCHA**

amb el títol

**"Los errores léxicos en textos escritos en español por  
alumnos universitarios brasileños en formación como  
profesores de Español Lengua Extranjera"**

Directora: María Rosa Torras i Cherta

per a l'obtenció del títol de Doctor/a en

**CIÈNCIES DE L'EDUCACIÓ**

Barcelona, 12 de juny de 2013

**Facultat de Formació del Professorat  
Departament de Didàctica de la Llengua i la Literatura**



UNIVERSITAT DE BARCELONA



## AGRADECIMIENTOS

Muchas personas han influido y colaborado en la realización de esta tesis.

Primeramente, quisiera expresar mi gratitud al Programa *Alban* por la beca (E04D027393BR) que me brindó la oportunidad de realizar los estudios de doctorado en la *Universitat de Barcelona*, así como a todas las personas e instituciones que hicieron posible esta investigación y la conclusión de este doctorado.

En la *Universidade Federal de Pernambuco* me gustaría agradecer a la directora del Centro de Artes y Comunicación, Dra. Virgínia Leal, por su apoyo y sus ánimos, así como a los profesores Dr. Miguel Espar, Dr. Alfredo Cordiviola, Dr. Vicente Masip y Dra. Maria Tereza Leal, por haberme acercado al fascinante mundo hispanohablante. Además, les agradezco inmensamente, como también al profesor Dr. Alberto Poza, por haber abierto sus aulas y permitido el acceso a los estudiantes. En especial, agradezco a todos los alumnos del curso de Letras sin cuya valiosa contribución esta investigación no habría sido posible. También expreso aquí mi gratitud al Dr. Rafael Dueire Lins por haberme dado la oportunidad de acercarme al mundo de la informática y a la investigación científica.

En la *Universitat de Barcelona* deseo agradecer fundamentalmente a mi tutora Dra. Rosa Maria Torras i Cherta por confiar en mí, por compartir sus conocimientos, por dedicarme su tiempo y también por su orientación objetiva, precisa y eficaz. A la Directora del Departamento de Didáctica de la Lengua y la Literatura durante mi formación Dra. Celia Romea Castro, sin cuyo apoyo no habría sido posible cumplir con los requisitos académicos y administrativos de los estudios a lo largo del curso. Al Coordinador del Doctorado y actual Director del Departamento Dr. Joan Perera, cuya orientación también fue fundamental en el cumplimiento de los requisitos del programa. A los profesores de las asignaturas de doctorado: Dra. Margarida Cambra, Dr. Joan Perera, Dr. Antonio Mendoza, Dr. Joan-Tomás Pujolà, Dra. Maria Rosa Torras, Dra. Celia Romea Castro y Dr. Miquel Llobera, muchas gracias por sus clases y por compartir su saber. También deseo expresar mi agradecimiento a Carme Núñez, administrativa del Departamento y a Teresa Fernández, administrativa de la Facultad, que siempre han atendido prontamente a las muchas peticiones, al profesor Jaume Llopis Pérez del Departamento Estadística, por enseñarme los conceptos básicos en estadística, esenciales en el desarrollo de esta investigación, y a la profesora María José Rubio del Departamento de Métodos de Investigación y Diagnóstico en Educación, por la ratificación de las pruebas estadísticas utilizadas en el análisis de los datos cuantitativos.

En la *University of Essex* le agradezco a Christina Gkonou por su amistad y por compartir sus ideas en el desarrollo de proyectos innovadores, y especialmente a la Dra. Beatriz de Paiva por la confianza depositada en mí y por compartir su experiencia en la enseñanza de la lengua portuguesa en el Reino Unido.

Finalmente, expreso mi gratitud a mis padres y a mi hermano por su cariño, por estar siempre presentes apoyándome en mis decisiones y por enseñarme los buenos valores que me han formado como persona. A mis tíos y primos por sus ánimos, y a mis abuelos su ejemplo. A mis suegros les agradezco por la motivación y el afecto demostrado a lo largo de los años. A Ricardo, gracias por todo. A los profesores y amigos Gláucya, Yluska, Teresa y en especial a Mizael e Ignacio, cuya colaboración fue fundamental en la recogida de datos del estudio exploratorio. A los amigos que nos acompañan desde siempre Adal, Clarissa y Roberto. A los queridos amigos María Eugenia, Mario, Eréndira, Yc, Rosa, Delphine, Guillermo, Conchi, Zailce, Lore y Guille, que conocimos al comenzar nuestra vida en Barcelona, gracias por ser nuestra familia allí y por compartir no sólo los buenos momentos sino también las transformaciones por las que pasamos todos durante el proceso de desarrollo del doctorado en esta ciudad encantadora que nos ha enamorado a todos. A Eréndira y a Yc, principalmente por los ánimos y la revisión estilística. A Rosa le agradezco inmensamente por su amistad, por compartir su saber en el desarrollo de proyectos académicos, por el apoyo logístico, por su acogida y sus cuidados a lo largo de los años y especialmente por la revisión lingüística del texto de esta tesis.

¡Muchas gracias!

# ÍNDICE

ÍNDICE DE FIGURAS .....	vii
ÍNDICE DE TABLAS .....	ix
Introducción.....	1
Justificación .....	5
Parte I: Marco teórico .....	9
1. Los paradigmas de estudio de la Adquisición de Segundas Lenguas...	10
1.1. El Análisis Contrastivo.....	13
1.2. El Análisis de Errores.....	16
1.3. La Interlengua.....	21
2. El aprendizaje de una segunda lengua y el conocimiento lingüístico previo.....	25
2.1. El estudio de la transferencia lingüística .....	25
2.1.1. Breve contextualización histórica.....	26
2.1.2. Variedad terminológica.....	29
2.1.3. Niveles lingüísticos en los que se verifica la transferencia.....	31
2.1.4. Condiciones que favorecen la transferencia .....	32
2.2. La transferencia lingüística y el léxico .....	34
2.3. La transferencia y la proximidad lingüística .....	40
2.4. La percepción de la proximidad lingüística .....	43
2.5. Lenguas afines: el portugués y el español.....	46
3. El error.....	55
3.1. La identificación, descripción y clasificación del error .....	55

3.1.1. Identificación del error.....	57
3.1.2. Descripción del error.....	58
3.1.3. Clasificación del error.....	59
3.1.4. Explicación del error .....	64
3.2. El error léxico.....	67
4. La fosilización y la estabilización .....	71
4.1. Las categorías de la fosilización .....	75
4.2. La fosilización y la influencia de la L1 .....	76
5. La producción escrita y las medidas de calidad .....	77
5.1. La producción escrita en el ámbito de las segundas lenguas .....	77
5.2. La fluidez .....	78
5.3. La diversidad léxica.....	79
5.4. La calidad del texto escrito .....	83
Parte II: El contexto general de la investigación.....	85
6. Contexto lingüístico .....	85
6.1. Las variedades del portugués .....	85
7. La enseñanza de español en Brasil.....	89
7.1. La enseñanza del español en Pernambuco .....	91
7.2. La selectividad: el <i>vestibular</i> .....	93
7.3. El curso de Letras de la <i>Universidade Federal de Pernambuco</i> (UFPE) .....	95
Parte III: Estudio empírico.....	97
8. Objetivos específicos y preguntas de investigación .....	97
9. Metodología .....	99

9.1. Los métodos de investigación cuantitativa en ASL.....	99
9.2. El diseño pseudolongitudinal.....	103
9.3. Participantes.....	105
9.4. Instrumento de recogida de datos lingüísticos .....	108
9.5. Procedimiento de recogida de datos .....	110
9.6. Organización del <i>corpus</i> lingüístico.....	112
9.7. Instrumentos para el procesamiento y análisis de los datos lingüísticos ...	114
9.7.1. El <i>Wordsmith Tools</i> y sus aplicaciones .....	115
9.7.2. Procesamiento estadístico (SPSS).....	120
Parte IV: Resultados, análisis y discusión .....	123
10. Resultados y análisis .....	123
10.1. Resultados y análisis de la fluidez y de la diversidad léxica .....	125
10.2. Resultados y análisis de errores léxicos y funcionales.....	130
10.3. Resultados y análisis de los errores según la clase de palabra léxica .....	134
10.4. Resultados y análisis del tipo de error léxico .....	143
10.5. Resultados y análisis del origen del error léxico.....	147
10.6. Resultados y análisis según el tipo y el origen del error .....	150
10.7. Resultados y análisis de la percepción de la proximidad lingüística .....	160
11. Discusión.....	165
11.1. La fluidez y la diversidad léxica .....	165
11.2. Los errores léxicos y funcionales.....	168
11.3. Los errores según la clase de palabra léxica.....	171
11.4. Los errores léxicos según el tipo .....	182
11.5. Los errores léxicos según el origen .....	196
11.6. Los errores léxicos según el tipo y el origen.....	200
11.7. La percepción de la proximidad lingüística .....	209

Parte V: Conclusiones y consideraciones finales .....	213
Conclusiones .....	213
Consideraciones finales .....	223
Bibliografía .....	229
Anexos .....	249
Anexo 1.....	251
Anexo 2.....	255
Anexo 3.....	257
Anexo 4.....	263
Anexo 5.....	265
Anexo 6.....	267
Anexo 7.....	273
Anexo 8.....	279



## ÍNDICE DE FIGURAS

Figura 1. Esquema conceptual de la tesis. ....	4
Figura 2. Los paradigmas de aprendizaje de L2/LE. ....	30
Figura 3. Esquema de clasificación de los errores de James (1998). ....	59
Figura 4. Medida de diversidad léxica Uber U. ....	82
Figura 5. Número de candidatos por lengua extranjera en el <i>vestibular</i> de la UFPE. ....	93
Figura 6. Número de candidatos al curso de Letras de la UFPE/nº de plazas. ....	94
Figura 7. Edad de los participantes de la muestra. ....	106
Figura 8. Fases de codificación de los datos con ejemplos. ....	116
Figura 9. Ejemplo de pantalla de la aplicación <i>Concord</i> . ....	118
Figura 10. Ejemplo de la aplicación <i>Wordlist</i> . ....	119
Figura 11. Medida de fluidez. ....	126
Figura 12. Medida de diversidad léxica: Uber U y Uber U sin errores. ....	128
Figura 13. Media del número errores léxicos y funcionales por texto. ....	132
Figura 14. Media de palabras léxicas y media de errores en las palabras léxicas por texto según la clase. ....	135
Figura 15. Media de palabras léxicas según la clase por texto. ....	137
Figura 16. Media de errores léxicos según la clase. ....	140
Figura 17. Media del tipo de error identificado por texto. ....	144
Figura 18. Media del número de errores según el origen en cada bloque de análisis. ....	148
Figura 19. Errores léxicos según el tipo y el origen. ....	151
Figura 20. Media de errores léxicos por texto según el origen y el tipo. ....	153
Figura 21. Media de errores interlingüísticos según el tipo. ....	154
Figura 22. Media de los errores intralingüísticos según el tipo por bloque de análisis. ....	156
Figura 23. Media de los errores ambiguos según el tipo por bloque de análisis. ....	158
Figura 24. Percepción de los participantes sobre la facilidad de escribir en español. ....	161
Figura 25. Razones de la dificultad para redactar los textos propuestos. ....	162
Figura 26. Autoevaluación de los participantes sobre las habilidades receptivas y productivas. ....	163
Figura 27. Autoevaluación de los participantes sobre la expresión escrita según el nivel. ..	164
Figura 28. Ejemplo de falsos cognados. ....	204



## ÍNDICE DE TABLAS

Tabla 1. Programa de lengua española del curso de Letras de la UFPE. ....	96
Tabla 2. Género de los participantes. ....	105
Tabla 3. Número de textos escritos obtenidos por asignatura en cada recogida de datos. .	111
Tabla 4. Recogida de datos y formación de bloques de análisis. ....	114
Tabla 5. Fases de codificación, códigos y número de palabras etiquetadas. ....	123
Tabla 6. Test U de Mann-Whitney: variable fluidez. ....	126
Tabla 7. Test U de Mann-Whitney: variable diversidad léxica (Uber U) ....	129
Tabla 8. Test U de Mann-Whitney: variable diversidad léxica (Uber U sin errores). ....	129
Tabla 9. Media de errores por texto. ....	131
Tabla 10. Ejemplos de errores léxicos y funcionales. ....	132
Tabla 11. Test U de Mann-Whitney: variable errores léxicos y funcionales. ....	133
Tabla 12. Media de palabras léxicas y media de errores en las palabras léxicas por texto según la clase. ....	135
Tabla 13. Ejemplos de errores léxicos según la clase de palabra. ....	136
Tabla 14. Test U de Mann-Whitney: variable clases de palabra léxica. ....	138
Tabla 15. Test U de Mann-Whitney, variable errores sustantivo. ....	141
Tabla 16. Test U de Mann-Whitney, variable errores adjetivo. ....	142
Tabla 17. Test U de Mann-Whitney, variable errores verbo. ....	142
Tabla 18. Test U de Mann-Whitney, variable errores adverbio. ....	142
Tabla 19. Ejemplos de errores léxicos según el tipo. ....	144
Tabla 20. Test U de Mann-Whitney: variable errores formales. ....	145
Tabla 21. Test U de Mann-Whitney: variable errores semánticos. ....	146
Tabla 22. Test U de Mann-Whitney: variable errores mixtos. ....	146
Tabla 23. Ejemplos de errores léxicos según el origen. ....	147
Tabla 24. Test U de Mann-Whitney: variable origen del error. ....	149
Tabla 25. Ejemplos de errores léxicos de origen interlingüístico según el tipo. ....	151
Tabla 26. Ejemplos de errores léxicos de origen intralingüístico según el tipo. ....	152
Tabla 27. Ejemplos de errores léxicos de origen ambiguo según el tipo. ....	152
Tabla 28. Test U de Mann-Whitney: variable error interlingüístico según el tipo. ....	155
Tabla 29. Test U de Mann-Whitney: variable tipo del error intralingüístico. ....	157

Tabla 30. Test U de Mann-Whitney: variable tipo del error ambiguo.....	159
Tabla 31. Test de normalidad: variables palabras total y errores total por asignatura. ....	267
Tabla 32. Test de normalidad: variable palabras total. ....	267
Tabla 33. Test de normalidad: variable errores.....	268
Tabla 34. Test de normalidad: variable categoría de palabra léxica. ....	269
Tabla 35. Test de normalidad: variable categoría de palabra léxica con errores.....	270
Tabla 36. Test de normalidad: variable tipo de error léxico.....	270
Tabla 37. Test de normalidad: variable origen del error léxico.....	271
Tabla 38. Test de normalidad: variable la riqueza léxica.....	271
Tabla 39. Test de Kruskal-Wallis: fluidez. ....	273
Tabla 40. Test de Kruskal-Wallis: diversidad léxica. ....	273
Tabla 41. Test de Kruskal-Wallis: errores. ....	274
Tabla 42. Test de Kruskal-Wallis: categoría de palabra léxica.....	275
Tabla 43. Test de Kruskal-Wallis: categoría de palabra léxica.....	275
Tabla 44. Test de Kruskal-Wallis: tipo del error.....	276
Tabla 45. Test de Kruskal-Wallis: origen del error.....	276
Tabla 46. Test de Kruskal-Wallis: error interlingüístico según el tipo. ....	277
Tabla 47. Test de Kruskal-Wallis: error intralingüístico según el tipo. ....	277
Tabla 48. Test de Kruskal-Wallis: error ambiguo según el tipo. ....	278

## INTRODUCCIÓN

El proceso de enseñanza-aprendizaje de una lengua próxima, en nuestro caso la enseñanza de la lengua española a estudiantes lusohablantes brasileños, supone facilidades y dificultades específicas relativas a varios aspectos del lenguaje. La experiencia muestra que en este contexto de aprendizaje, de manera general, hay comunicación en la lengua meta a partir del primer día de clase desde un nivel inicial, pues los alumnos comprenden las instrucciones del profesor y responden/actúan de forma adecuada. Sin embargo, percibimos que esta facilidad en la comprensión tanto oral como escrita no permite, muchas veces, que el alumno avance en el aprendizaje del español. Asimismo, se notan rasgos de elementos que persisten en su interlengua durante este período.

Motivados por conocer los errores propios de la expresión escrita en la producción lingüística de los estudiantes brasileños de español, desarrollamos un estudio exploratorio con una muestra reducida de participantes que se presentó en el año 2006 en el Departamento de Didáctica de la Lengua y la Literatura de la Universidad de Barcelona con el fin de superar el examen de suficiencia investigadora (DEA - Diploma de Estudios Avanzados). En ese estudio centramos nuestra atención en los errores léxicos, porque por un lado es en las palabras léxicas (sustantivo, adjetivo, verbo, adverbio) que se encuentra la carga semántica del mensaje, y por otro, los errores léxicos son los que suelen causar interrupción en la comunicación con hablantes nativos del idioma. Tras el análisis de los textos, se identificaron los errores en las palabras léxicas, el tipo de error (formal o semántico) y su origen (si se debía a la influencia de la lengua materna o no).

Una vez obtenida la suficiencia investigadora y partiendo de ese estudio previo, planteamos una investigación más completa y profunda que da forma a esta tesis doctoral. Se trata de un estudio con una muestra más amplia de estudiantes brasileños en diferentes niveles del período de formación como profesores de español en el ámbito universitario, cuyo objetivo general es analizar los errores léxicos en la producción escrita para identificar su presencia, su frecuencia y su origen en diferentes niveles del programa de formación, y, además, explorar una posible tendencia a la persistencia que puedan presentar dichos errores.

La tesis se organiza como se muestra en el esquema que se presenta a continuación en la figura 1. Para empezar, en la justificación presentamos las razones que han motivado el desarrollo de esta investigación y continuación, explicamos la organización de la tesis que se divide en cuatro partes. En la primera parte, se presenta el marco teórico de la literatura que fundamenta la presente investigación. En el capítulo 1, se hace una exposición general de los paradigmas de estudio de la adquisición de segundas lenguas (ASL) desde cuando surgió el Análisis Contrastivo (AC), pasando por el Análisis de Errores (AE), hasta llegar a la hipótesis de la Interlengua (IL). En el capítulo 2, se abordan las teorías de transferencia lingüística, interferencia e influencia croslingüística (ICL). También se comentan los resultados de algunos trabajos que estudiaron la transferencia específicamente en el ámbito del léxico y las implicaciones de la proximidad lingüística en los fenómenos de transferencia. En este capítulo también se aborda el tema de la percepción que tienen los estudiantes con respecto a la proximidad tipológica y las características del aprendizaje de una lengua en un contexto de lenguas afines como son el portugués y el español. En el capítulo 3, se repasa el concepto de error como elemento positivo inherente al proceso de aprendizaje, se describen los criterios para la identificación, descripción y clasificación del error. También se comentan las posibles causas de los errores y, además, se dedica especial atención a los errores léxicos, objeto de esta investigación. El capítulo 4 se centra en el tema de la fosilización y la estabilización. En este capítulo se explican las categorías de la fosilización, su relación con el factor edad y también con la primera lengua (L1) o lengua materna. El capítulo 5, que finaliza la primera parte, trata de la producción escrita y las medidas de calidad como son la fluidez, la diversidad léxica y la corrección. Asimismo, se comentan los métodos más utilizados que se han desarrollado con el objetivo de medir la calidad del texto escrito.

En la segunda parte se describe el contexto general donde se ha llevado a cabo la investigación. En el capítulo 6 se explica el contexto lingüístico y en el capítulo 7 se ofrece una visión de la situación actual de la enseñanza del español en Brasil en la enseñanza secundaria y media. Además se hace una breve explicación de las características y de la importancia de los exámenes de selectividad, conocidos en Brasil como '*vestibular*', para la enseñanza de lenguas extranjeras. A continuación se describen las características del curso

de Letras de la *Universidade Federal de Pernambuco* (UFPE), donde se ha realizado la recogida de datos.

En la tercera parte presentamos el estudio empírico. En el capítulo 8 se detallan los objetivos específicos y se definen las preguntas de esta investigación. En el capítulo 9, que trata de la metodología, se describen a los participantes de la muestra y sus características, los instrumentos y el procedimiento utilizado en el proceso de recogida de datos, la organización del *corpus* lingüístico en tres bloques para facilitar el análisis de los datos, así como los instrumentos utilizados para el procesamiento y análisis de los datos.

La cuarta parte de este trabajo comprende la presentación de los resultados y análisis en el capítulo 10 y la discusión de los resultados en el capítulo 11. Finalmente, presentamos las conclusiones y las consideraciones finales en las que sugerimos algunas implicaciones didácticas derivadas de la reflexión realizada a lo largo del desarrollo de esta investigación y líneas de trabajos futuros, así como la bibliografía consultada en la que nos apoyamos. Al final, se adjuntan los anexos a los que hacemos referencia a lo largo del cuerpo de la tesis, así como el *corpus* lingüístico en formato digital (CD).

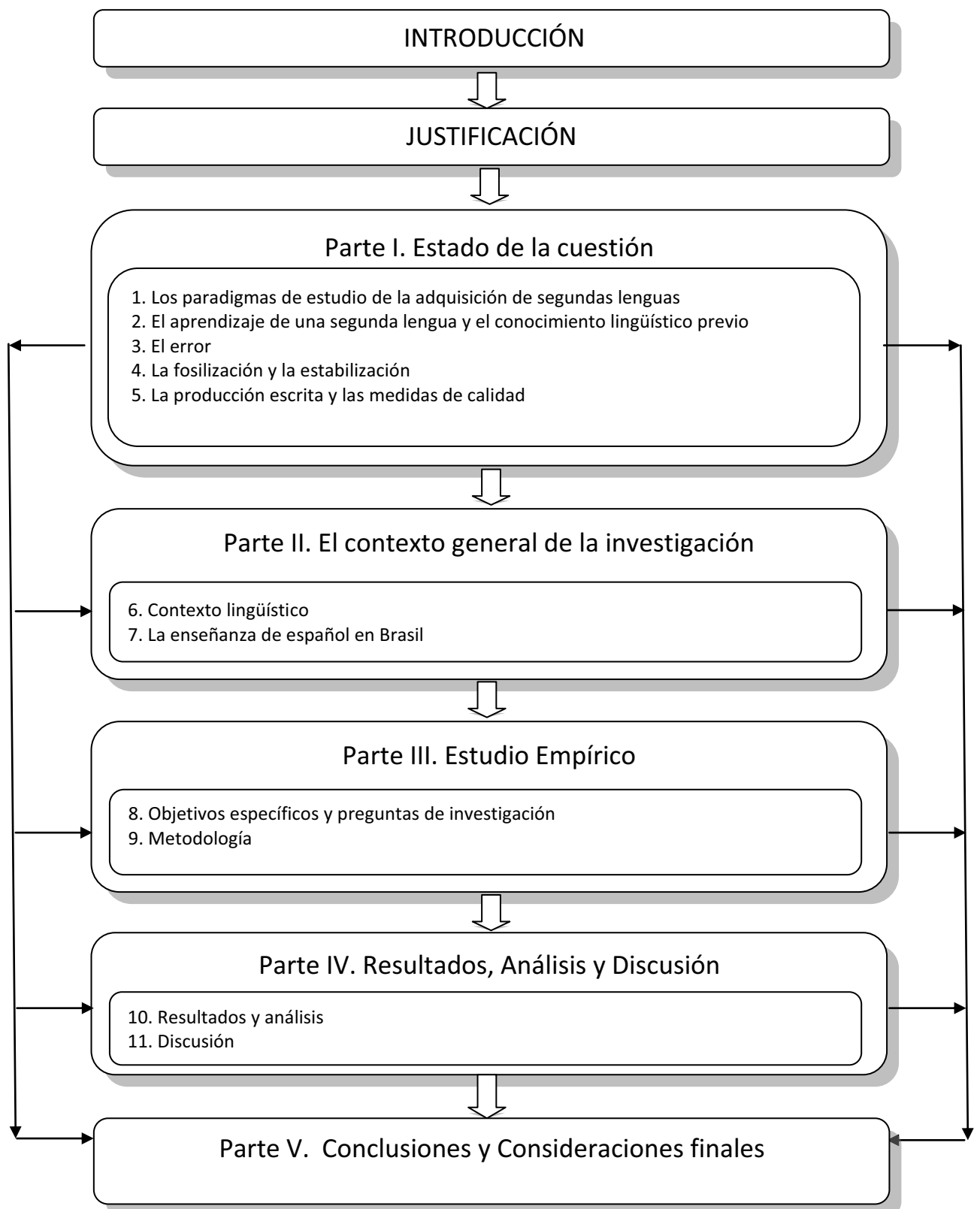


Figura 1. Esquema conceptual de la tesis.



## JUSTIFICACIÓN

Esta investigación surge del interés por la enseñanza de lenguas y de la práctica como profesora de español que empezó en 1999 en Brasil, en academias de lengua y en el Departamento de Letras de la *Universidade Federal de Pernambuco* (UFPE) y que prosigue hasta hoy en Inglaterra, como profesora de portugués en la *University of Essex* y de portugués y español en la *Regents' University London*. La doble experiencia, como alumna y como profesora del curso de Letras de la UFPE, ha permitido reflexionar sobre los más variados aspectos involucrados en la formación lingüística de futuros profesores de lengua española en un contexto de lenguas afines y ha posibilitado hacer una apreciación más detallada de las características del contexto en el que centramos esta tesis. Posteriormente, la vivencia como profesora de portugués y español en Inglaterra, un contexto en el que las lenguas implicadas son tipológicamente distintas salvo en casos aislados en los que el alumno tiene algún conocimiento de una lengua afín, ha fortalecido la apreciación que hacemos en esta tesis sobre la importancia de tener en cuenta la especificidad propia de cada contexto de enseñanza-aprendizaje para promocionar un aprendizaje más eficaz.

Como es sabido entre la mayoría de los profesores de español en Brasil, la proximidad y semejanza entre las lenguas española y portuguesa posibilita comunicación, principalmente oral, pero también escrita, debido a que ambas lenguas comparten gran parte del léxico. La doble experiencia vivida como alumna y docente de esta lengua, y también como hablante del portugués variedad brasileña (L1), nos ha mostrado que la transparencia lingüística ofrece una facilidad inicial en el contacto con la lengua española, que se convierte a lo largo del tiempo en una dificultad para adquirir un nivel de corrección lingüística alto con respecto a determinadas características propias de la lengua española, tanto sintácticas como aspectos de precisión léxica.

Es por ello que hemos planteado este estudio con alumnos brasileños de español, más concretamente con los alumnos de español del curso de Letras de la UPFE, porque además de ser estudiantes de lengua española, se están formando como profesores de ELE, y como profesores es esencial que puedan ofrecer un modelo de lengua libre de errores. La detección de los errores léxicos persistentes en la interlengua, y consecuentemente potencialmente fosilizables, puede ser de gran ayuda no sólo para los futuros profesores de

lengua española sino para todos aquellos aprendices de habla portuguesa que quieran alcanzar un dominio correcto de la lengua española. Creemos que este conocimiento puede tener una aplicación práctica en el diseño curricular tanto a nivel didáctico y programático como en el diseño de materiales didácticos.

Al empezar a investigar sobre la adquisición del español como lengua extranjera, se ha encontrado una gran variedad de estudios relacionados con el aprendizaje de los aspectos morfosintácticos de la lengua. También es cierto que al reflexionar sobre nuestra experiencia como alumna y docente de español, es notable el énfasis que se suele dar a los aspectos sintácticos de la lengua meta, quizá por ser éste un sistema cerrado al cual se accede fácilmente a través de las gramáticas y libros didácticos disponibles. Con todo, el interés de esta investigación se ha centrado en el estudio de los errores léxicos porque, en primer lugar, están en juego dos lenguas semejantes con elevado número de palabras cognadas por lo que nos interesa saber de qué manera esta característica interfiere en el aprendizaje. En segundo lugar, porque es en el léxico que se encuentra la carga semántica del mensaje y los fallos en este aspecto son los que más suelen obstruir o interrumpir la comunicación.

Se ha elegido analizar la producción escrita porque ésta es una actividad que hace parte de la vida académica de los estudiantes y no representa una situación nueva para ellos, porque a lo largo del curso de Letras los alumnos tienen que escribir trabajos, ensayos literarios, exámenes, informes de observación de clase, diarios, planes de clase, etc. Además, el hecho de escribir significa también que el alumno dispone de algún tiempo para ordenar su pensamiento y tomar decisiones sobre la elección de vocabulario y elaboración de estructuras, por lo que pensamos que los desvíos que puedan presentar son ya el resultado de una reflexión previa. Asimismo, queremos hacer hincapié en el hecho de que consideramos el desarrollo de las cuatro destrezas comunicativas (hablar, escuchar, leer y escribir) igual de importantes en el proceso de enseñanza-aprendizaje de lenguas. El hecho de que nos centremos en el estudio de los errores léxicos en textos escritos no significa que demos más importancia a esta destreza, simplemente delimitamos el aspecto que nos interesa investigar en la presente investigación.

Partimos de la idea de que el estudio de la lengua que los estudiantes producen a lo largo de su proceso de adquisición-aprendizaje constituye la principal fuente de información para profundizar en la investigación en adquisición de segundas lenguas. Nuestro objetivo principal es conocer las características de los textos escritos en español por alumnos brasileños, más concretamente los aspectos referentes a los errores léxicos. Nos interesa averiguar si hay y de qué tipo son los errores de vocabulario, si se deben a interferencias de su L1, el portugués variedad brasileña, de otras lenguas que conocen o de su propia interlengua y si éstos se dan a nivel semántico u ortográfico. También nos interesa averiguar cuáles son los errores léxicos que persisten en los niveles más altos.

Hemos notado un aumento en el número de investigaciones dedicadas específicamente a la enseñanza y al aprendizaje de la lengua española en Brasil, probablemente impulsadas por la inclusión de esta lengua en el currículum obligatorio de la enseñanza secundaria a partir del año 2010. Es nuestro deseo que este estudio no solo contribuya a un mayor conocimiento de la adquisición de una lengua en un contexto de lenguas próximas, sino que también pueda aportar directrices para mejorar la enseñanza del español a hablantes de portugués.



## PARTE I: MARCO TEÓRICO

Esta tesis se enmarca en el ámbito de la investigación en Adquisición de Segundas Lenguas (ASL), término que se utiliza de manera general para referirse al campo de investigación de la enseñanza-aprendizaje y de la adquisición de una lengua después de que la lengua materna o primera lengua (L1) del estudiante haya sido aprendida (Gass y Selinker, 2008).

Se parte de la idea de que el estudio de la lengua que los estudiantes producen a lo largo de su proceso de adquisición-aprendizaje es una manera de profundizar en el conocimiento que se tiene sobre el área. Ese pensamiento ha guiado el diseño de la presente investigación, que tiene como objetivo estudiar los errores léxicos en producciones escritas de estudiantes lusohablantes brasileños durante su período de formación como profesores de español como lengua extranjera (ELE).

La enseñanza de la lengua española en Brasil se da en un contexto de lengua extranjera (LE), es decir, un entorno en el que la lengua objeto de estudio no pertenece al país donde se estudia y el contacto de los estudiantes con esta lengua se da casi exclusivamente en el aula. Aún así, en este trabajo se utilizan los términos segunda lengua (L2) y lengua extranjera (LE) indistintamente, porque, así como lo hace Ellis (2008), consideramos como segunda lengua cualquier lengua estudiada o aprendida por el individuo después de que éste haya adquirido su lengua materna (L1).

A continuación se presentan los capítulos que componen esta primera parte del trabajo y los supuestos teóricos en los que se apoyan esta investigación.

## 1. LOS PARADIGMAS DE ESTUDIO DE LA ADQUISICIÓN DE SEGUNDAS LENGUAS

El estudio de la Adquisición de Segundas Lenguas (ASL) se estableció como campo de investigación en los años sesenta. Según Ellis (2008), se tomó los sesenta como el inicio debido al impacto que los trabajos desarrollados en esos años, motivados en gran parte por el deseo de mejorar la pedagogía utilizada en la enseñanza de segundas lenguas, tuvieron en la investigación actual.

El objetivo de las investigaciones en ASL es conocer el proceso que subyace al aprendizaje de una lengua no nativa después del aprendizaje de la lengua materna o primera (L1). La lengua no nativa se conoce como segunda lengua (L2), independientemente de si es la tercera, cuarta o quinta lengua que aprende el individuo. Los estudios en ASL se dedican a investigar cómo los estudiantes crean un nuevo sistema lingüístico a través de la exposición limitada a la L2 y también qué aspectos de una L2 se aprenden y cuáles no. Además, interesa a los investigadores conocer por qué una gran parte de los estudiantes de L2 no llega al mismo nivel de conocimiento y competencia en L2 como lo hacen en su L1 y de por qué sólo algunos estudiantes parecen conseguir una competencia como la de un nativo en más de una lengua. También es de interés de los estudiosos investigar la naturaleza de las hipótesis que los alumnos tienen sobre las reglas de la L2 (Gass y Selinker, 2008).

El aprendizaje de una segunda lengua puede darse en contextos variados, como son los contextos de instrucción formal, cuando el estudiante frecuenta un curso en una escuela o academia, o en un contexto natural, en el que el individuo está en exposición a la lengua en la comunidad en la que se habla. Se denomina lengua extranjera (LE) la lengua que no se utiliza como medio de comunicación por los miembros de la comunidad y que suele estudiarse en un contexto formal de enseñanza. Por otro lado, la lengua que se aprende en un contexto natural de inmersión lingüística porque se utiliza como medio de comunicación por los miembros de la comunidad, se conoce como segunda lengua (L2), es decir, la L2 “cumple una función social e institucional en la comunidad lingüística en que se aprende” (Santos Gargallo, 1999: 21). Se suele diferenciar el aprendizaje de LE del aprendizaje de L2, principalmente porque el primero hace referencia a la lengua no nativa dentro de un contexto de lengua nativa, por ejemplo, hablantes de portugués aprendiendo español en Brasil, generalmente en el aula. El segundo fenómeno hace referencia al aprendizaje de una

lengua no nativa en un contexto en el que se habla la lengua meta, por ejemplo, hablantes de portugués aprendiendo español en España.

Se ha observado que el contexto en el que se desarrolla el aprendizaje de una L2 repercute en la manera como el aprendiente asimila la lengua. Es por ello que hay autores como Krashen (1981) que diferencian los términos adquisición, cuando la asimilación de una lengua ocurre a través de la exposición a un *input*<sup>1</sup> lingüístico comprensible en la L2 en contextos naturales, y aprendizaje, cuando la asimilación de una L2/LE se da en un contexto formal de instrucción. Para estos investigadores, la adquisición es el proceso subconsciente en el que el hablante tiene que estar expuesto a muestras de lengua para adquirirla y aprendizaje se refiere al proceso consciente de estudio de la lengua. Sin embargo, Ellis (2008) afirma que aunque esta distinción sea válida, es muy difícil demostrar cuál es el proceso implicado en cada momento de la formación del estudiante y por eso él opta por emplear los términos adquisición y aprendizaje indistintamente. Asimismo, Llopis García (2009) afirma que las fronteras entre el aprendizaje y la adquisición son imprecisas y que la ASL engloba todas las situaciones, incluidas las que combinan contextos de adquisición en un ambiente de LE o aprendizaje en situaciones de inmersión (L2). El aprendizaje tanto de una LE como de una L2, como se ha comentado anteriormente, puede realizarse mediante instrucción formal o puede ocurrir de manera natural. La principal diferencia es que el aprendizaje en un ambiente de segundas lenguas puede proporcionar un mayor acceso de los estudiantes a la lengua meta (L2) en comparación con el aprendizaje en el contexto de lenguas extranjeras (Gass y Selinker, 2008).

Los estudios realizados en el campo de la ASL tienen, muchas veces, implicaciones para la enseñanza de lenguas. Estos estudios han colaborado para que los profesores de lenguas y diseñadores de currículos sean conscientes de que el aprendizaje de lenguas consiste en algo más que la memorización de reglas, es decir, en fomentar el desarrollo de la competencia comunicativa del estudiante. Esa nueva conceptualización en aprendizaje

---

<sup>1</sup> Según el diccionario de términos clave de ELE del Instituto Cervantes (Centro Virtual Cervantes, 2012), el término inglés *input* es también conocido como “caudal lingüístico” o “aducto”, y se usa para hacer referencia a las muestras de lengua meta a las que el estudiante está expuesto en situaciones comunicativas durante su proceso de aprendizaje. Sin embargo, el alumno sólo es capaz de comprender una parte de esta información, y por consiguiente, retenerla en su memoria a corto plazo (*intake* o apropiación de datos).

lingüístico ha llevado a la creación de métodos<sup>2</sup> de enseñanza que enfatizan la comunicación. Ese es el caso del enfoque comunicativo (*Communicative Approach*) y del enfoque por tareas (*Task-Based Approach*) que procuran trabajar aspectos pragmáticos en la enseñanza de lenguas. Como bien explica Agudelo (2011: 66) “el objetivo didáctico del enfoque comunicativo – y del enfoque por tareas – es el desarrollo de la competencia comunicativa, o la capacidad de interaccionar lingüísticamente de forma adecuada en situaciones de comunicación, integrando para ello las habilidades necesarias”.

Otro aspecto importante de la ASL relacionado con la pedagogía de lenguas es la expectativa que los profesores tienen de los alumnos. Un ejemplo de ello es el cambio en el concepto de error que ha dejado de ser visto como algo negativo que debería ser evitado y ha pasado a ser concebido como parte inherente del proceso de enseñanza-aprendizaje y producto de las inferencias que hace el alumno en el proceso de construcción de la gramática de la lengua meta. El error revela información relevante sobre las hipótesis que tiene el estudiante sobre la L2 y cuanto más los profesores conozcan sobre el proceso de aprendizaje, más pueden ayudar a sus alumnos en la construcción del conocimiento lingüístico.

De acuerdo con Gass y Selinker (2008), conocer bien una lengua significa tener un conocimiento similar al de un nativo sobre su propia lengua. Dado que de manera general, el conocimiento lingüístico engloba la fonología, la sintaxis, la morfología, el léxico, la semántica y la pragmática, el hablante no nativo tiene que conocer y desarrollar todos estos aspectos lingüísticos en su proceso de aprendizaje de la L2.

Actualmente, el supuesto básico de la ASL es el de que los estudiantes de una L2 crean un sistema lingüístico propio conocido como interlengua (IL). La interlengua se caracteriza como un sistema inestable que tiene elementos de la L1 del estudiante, de la lengua que está estudiando (L2) y de elementos que no tienen relación con la L1 ni con la lengua meta (L2), que Gass y Selinker (*op. cit.*) han llamado ‘formas nuevas’ (*new forms*). Un aspecto central en el concepto de interlengua es el de la *fosilización* (se profundiza en este tema en el capítulo 4), que se refiere a la interrupción del aprendizaje, generalmente observada a

---

<sup>2</sup> Para una revisión de los diferentes métodos de enseñanza utilizados en la enseñanza de lenguas a lo largo del tiempo, se recomienda la consulta de la tesis de Agudelo (2011) que se centra en la enseñanza comunicativa de ELE.



través de errores persistentes que se detectan en la interlengua. Esa interrupción puede darse de forma global o afectar solo a determinados aspectos lingüísticos. Es importante destacar que Long (2003) considera el término 'estabilización' más apropiado para referirse al fenómeno de la fosilización, porque él advierte de que es muy difícil determinar cuándo el aprendizaje de una lengua ha terminado de ocurrir del todo.

Dado que este trabajo se centra en el estudio del error, se repasan, a continuación, las características de los principios teóricos y metodológicos de los tres paradigmas de estudio de los errores en la enseñanza de segundas lenguas: el Análisis Contrastivo (AC), el Análisis de Errores (AE) y la Interlengua (IL). Además, se comentan las transformaciones que dichos paradigmas han sufrido a lo largo del tiempo a medida que surgían nuevas investigaciones en el área.

### 1.1. El Análisis Contrastivo

En los años cincuenta y sesenta, el paradigma de estudio del aprendizaje de segundas lenguas (L2) era el Análisis Contrastivo (AC), que surgió como un intento de encontrar explicaciones teóricas para los errores que los alumnos cometían cuando producían L2, porque se suponía que los estudiantes transferían rasgos de su L1 a la lengua que estaban aprendiendo. Este modelo se basaba en la comparación tanto de la forma como del significado de las palabras de las dos lenguas relacionadas en el proceso: la lengua materna (L1) del estudiante y la L2, la lengua meta.

Para Lado (1957), el error no era nada más que el resultado de la transferencia de hábitos de la L1, que podrían resultar en transferencia positiva o transferencia negativa. La transferencia positiva ocurría cuando el alumno utilizaba estructuras de la lengua materna (L1) similares a las estructuras de la L2, y la transferencia negativa (interferencia) se daba cuando el alumno hacía uso de alguna estructura de la L1 que no correspondía a la estructura de la lengua meta (L2). Dependiendo de la proximidad entre la L1 y la L2, se podía explicar o prever una parte de los errores que los alumnos producían como resultado de la transferencia equivocada del sistema de la lengua materna a la lengua meta e intentar adecuar la enseñanza para evitarlos. Se creía que la probabilidad de producirse errores en la

L2 estaba directamente relacionada con la distancia lingüística de las estructuras entre las lenguas implicadas.

Sin embargo, Gass y Selinker (2008) hacen hincapié en la caracterización de la transferencia como proceso y afirman que el resultado de este proceso de transferencia puede ser tanto positivo, cuando los elementos transferidos coinciden con los de la lengua meta, o puede ser negativo, cuando se produce el error, una vez que los elementos transferidos no coinciden con los de la lengua meta. Los autores destacan, por lo tanto, que no existen dos procesos de transferencia, uno positivo y otro negativo, sino que la transferencia es un proceso único y el resultado de la transferencia se verificaría una vez analizada la producción lingüística del estudiante.

Para el Análisis Contrastivo, paradigma de estudio que se basaba en la lingüística estructural y en el conductismo, interesaba la descripción de la lengua y la predicción del error. La principal causa del error era la interferencia de la L1 (transferencia negativa) y el error se consideraba una transgresión del sistema lingüístico de la L2 y que debía, por lo tanto, evitarse. Los esfuerzos pedagógicos se centraban en la enseñanza de las diferencias entre las lenguas en el intento de impedir la transferencia negativa de la L1 en la lengua meta.

El Análisis Contrastivo fue muy utilizado con el método audiolingüístico<sup>3</sup> (*Audio Lingual Method*), en conjunto con teorías lingüísticas estructuralistas y de psicología conductista. La idea de usar el AC en la enseñanza se basaba en el hecho de que el alumno, de forma consciente, pudiera traspasar los elementos de su L1 a los casos en que se podía aplicar en la L2, o sea, se intentaba crear el hábito de transferir los elementos comunes de la lengua materna a la lengua meta.

En el marco de la hipótesis del Análisis Contrastivo había dos versiones. Una de ellas era llamada *a priori* y la otra, *a posteriori*. La hipótesis del análisis contrastivo *a priori*, también era conocida como 'versión fuerte' (*strong version*) y la llamada *a posteriori* era también

---

<sup>3</sup> De acuerdo con Zanón (2007), el término original "*Audiolingualism*" fue creado por Nelson Brooks a finales de los años cincuenta. Según el Diccionario de Términos Clave de ELE, el método audiolingüe, conocido también como método audiolingüístico, pretende que el estudiante consiga obtener un dominio oral de la L2, automatizando el habla de manera que no tenga que recurrir a la traducción de la lengua materna. Una crítica a este método es la de que el estudiante tiene un papel limitado en el aprendizaje, dado que se espera que responda a los estímulos ofrecidos por el profesor sin considerar el contenido de lo que practica (Centro Virtual Cervantes, 2012).

conocida como 'versión débil' (*weak version*) (Wardhaugh, 1970). En la versión fuerte, *a priori*, como su nombre indica, se defendía la idea de que se podía hacer predicciones sobre el aprendizaje con base en la comparación entre las lenguas del aprendiz (L1) y la lengua meta (L2). En el caso de la versión débil, *a posteriori*, se analizaban los errores recurrentes de los alumnos y sólo en un segundo momento se comparaban esos errores con las diferencias entre la L1 y la L2 (Gass y Selinker, 2008).

Con el tiempo, las investigaciones empezaron a ratificar que a través de la versión fuerte del Análisis Contrastivo no se podía explicar la mayoría de los errores, ya que sólo una parte de ellos ocurría debido a la interferencia de la lengua materna. Además, las predicciones no siempre se confirmaban al analizarse la producción de los alumnos.

A principios de los años setenta, esta teoría pasó a ser cuestionada por su asociación con un modelo ultrapasado de descripción de lengua, el estructuralismo, y una teoría de aprendizaje desacreditada, el conductismo, que era la base del método de enseñanza audiolingüístico (James, 1998).

Los estudios de adquisición de L1 que afirmaban que los hablantes nativos aprendían la lengua construyendo reglas mentales contradecían la idea de que los hablantes aprendían su L1 como una serie de hábitos, lo que despertó el interés por conocer mejor la lengua que los estudiantes de L2 producían (Ellis, 2008). A partir de entonces, investigadores como Corder (1967, 1981) pasaron a centrar más su atención en la producción de los estudiantes y no tanto en la comparación de los sistemas lingüísticos de las L1 y L2. El Análisis Contrastivo (AC) fue, por lo tanto, dando lugar al Análisis de Errores (AE) debido a la constatación de que el AC ya no atendía a las necesidades de los investigadores.

Además, Odlin (2006: 26) afirma que para que se realice un análisis contrastivo completo, es necesario tener en cuenta la transferabilidad (*transferability*), es decir, la probabilidad con la que una determinada estructura puede transferirse de la L1 a la L2 (Kellerman, 1983). Éste es un concepto que lleva en consideración la probabilidad de factores como la naturaleza de la estructura transferida, la distancia entre las lenguas, la diferenciación (*markedness*) entre otros, y que ésta sólo es satisfactoria si considera los factores afectivos relevantes. Según Odlin (1992), los tres principios que pueden ayudar a identificar efectos de la transferibilidad en la situación de lenguas en contacto son el rango similar de distribución de

L1 y L2, las ocurrencias geográficas múltiples y la alta probabilidad en ciertas áreas geográficas.

Sin embargo, Jordan (1991) defiende la idea de que a pesar de presentar serios fallos con respecto al contraste de lenguas tipológicamente distintas, el Análisis Contrastivo es beneficioso cuando se aplica a la enseñanza de lenguas próximas a la L1 del alumno, o a otra L2 que haya estudiado anteriormente. La autora está de acuerdo con las críticas que se han hecho sobre el AC, pero defiende su utilización con fines pedagógicos en la enseñanza de portugués para hablantes de español, por ejemplo, porque son lenguas con un sistema lingüístico similar.

## 1.2. El Análisis de Errores

El Análisis de Errores (AE) se basa en la lingüística generativo-transformacional y en el cognitivismo, y su interés está en contribuir en el ámbito de la Adquisición de Segundas Lenguas (ASL). El Análisis de Errores predicaba el estudio de los errores producidos por los alumnos y los comparaba con la lengua meta (L2), en lugar de la comparación entre sistemas lingüísticos, como se hacía en el Análisis Contrastivo (AC). Los investigadores buscaban describir y explicar los errores a partir del estudio del sistema lingüístico producido por los alumnos para poder encontrar sus causas y entender las estrategias que los alumnos utilizaban en el proceso de aprendizaje de una lengua extranjera, es decir, a través del estudio de los errores procuraban conocer los mecanismos psicolingüísticos que se activan para compensar las deficiencias del estadio en el que se encuentra el sistema lingüístico del estudiante. El AE guarda cierta similitud con la versión débil del AC en lo que se refiere al análisis de los errores de los alumnos. Sin embargo, la diferencia está en que en el AE, como se ha dicho anteriormente, los errores se comparan con la lengua meta (L2) (Gass y Selinker, 2008).

Por lo tanto, en el AE se defendía la idea de que se podía describir los errores con base sólo en la lengua meta sin la necesidad de referirse a la L1 de los alumnos, y que, a diferencia del AC, no se tomaba en consideración la L1 de los estudiantes. Corder (1967) fue uno de los primeros investigadores en percibir que los errores revelaban indicios de cómo se aprende una segunda lengua. Según ese autor, los errores de los estudiantes indican tanto su

competencia como su modo de aprendizaje. Esta preocupación llevó a importantes cambios en las teorías lingüísticas de aprendizaje por su carácter didáctico (James, 1998).

Uno de los objetivos de un análisis de errores es establecer un inventario de los errores más frecuentes, valorando la importancia y gravedad de los mismos, con el objetivo de señalar las áreas de dificultad en el aprendizaje de una L2 para un determinado grupo de estudiantes con idéntica L1 (Santos Gargallo, 1993). No obstante, el AE no ha tenido en cuenta la variabilidad en la lengua del estudiante, es decir, si el alumno produce una palabra correcta en un determinado contexto y luego utiliza esa misma palabra, pero de manera equivocada, en otro contexto, es muy difícil decir si el alumno conoce o no la palabra y si su producción equivocada es un error o una falta (Ellis, 2008). Lo cierto es que los errores proporcionan evidencia del estadio del conocimiento que tiene el estudiante de la L2.

Al utilizar el AE los investigadores se dedicaron a observar las producciones de los alumnos y así se dieron cuenta de que el proceso de aprendizaje de lenguas era más complejo de lo que habían creído hasta el momento. Así como en la adquisición de la L1 por niños, se ha observado que, en el ámbito de la enseñanza de segundas lenguas, los errores no son el producto de una imitación errónea, pero sí evidencia de que existe un sistema de reglas subyacente (Gass y Selinker, *op. cit.*).

Para que un análisis de errores produzca un *corpus* con un nivel de fiabilidad lo más alto posible es recomendable seguir algunos pasos fundamentales (Santos Gargallo, *op. cit.*: 98): a) determinar el aspecto que se quiere analizar; b) diseñar el perfil del alumno; c) elegir el modelo de prueba más apropiado para satisfacer sus objetivos; d) determinar las instancias que se consideran erróneas y describirlas; e) explicarlas y valorar su gravedad de acuerdo con los criterios establecidos; f) confrontar su diagnóstico con el de otros colegas o, si es posible, con los propios informantes; g) generalizar.

Como se ha dicho más arriba, el método utilizado para estudiar el error consistía en tomar las producciones de los alumnos, y para ello se seguían los siguientes pasos: recogida de datos; identificación de los errores en su contexto; clasificación y descripción de los errores; cuantificación de los errores; explicación del error, averiguando los mecanismos o estrategias psicolingüísticas utilizadas e identificando las fuentes de cada error; si el análisis

tiene objetivos didácticos, se evalúa la gravedad del error y se busca la posible terapia con base en el tipo y frecuencia del error (Corder, 1967).

De esa corriente se destaca la importancia que genera en la nueva concepción del error, visto ahora como un paso relevante que hace parte del proceso de adquisición-aprendizaje de una lengua y que revela una competencia transitoria. Didácticamente, esta contribución permite la consideración positiva del error y consecuentemente la pérdida del miedo al error, lo que conlleva a que los alumnos produzcan más e incrementen las oportunidades de aprendizaje (Fernández, 1997).

Además, se ha verificado que el Análisis de Errores ofrece más posibilidades que el Análisis Contrastivo a la hora de explicar el error porque a diferencia del AC que se centraba sólo en la L1 como causa de los errores, el AE contempla por lo menos dos causas del error: una de origen interlingüístico, en la que el error se debía a la influencia de la L1 del hablante, y otra de origen intralingüístico, en la que el error se debía al propio proceso de enseñanza-aprendizaje de la L2, es decir, de desarrollo de la interlengua (Gass y Selinker, 2008).

Ellis (2008) apunta que el Análisis de Errores tiene una serie de limitaciones metodológicas y teóricas. Según este autor, al enfocar solamente en los errores que producen los estudiantes en un determinado punto en el tiempo, como hicieron la mayoría de estos estudios, sólo se ofrece una fotografía parcial, ya que no toma en consideración lo que el alumno hace correctamente, no se observa su desarrollo a lo largo del tiempo, ni se da cuenta del fenómeno de la *evitación*<sup>4</sup> (*avoidance*). En este sentido, Gass y Selinker (*op. cit.*) también señalan que el hecho de que los estudios realizados en el ámbito del Análisis de Errores no tuviesen en cuenta otros factores, como por ejemplo, aquello que los estudiantes hacían correctamente, era un obstáculo que impedía que los investigadores pudiesen tener una visión general del comportamiento lingüístico del estudiante.

El artículo de Schachter (1974) en el que se estudia la adquisición de las oraciones de relativo en inglés por hablantes de chino, japonés, árabe y persa es un ejemplo de cómo el

---

<sup>4</sup> La "evitación" es una estrategia de los estudiantes de segundas lenguas en la que se evita el empleo de determinadas formas o estructuras lingüísticas, lo que indica que tienen una dificultad con dicha forma y por eso no la utilizan en su discurso (Schachter, 1974). Además, Laufer (2000) destaca el hecho de que la utilización de la estrategia de la evitación implica el conocimiento de la forma o estructura que no se ha utilizado, una vez que no se puede evitar lo que no se conoce.

hecho de no tener en cuenta el todo puede enturbiar los resultados de una investigación. La autora argumenta que los estudiantes hablantes de chino y japonés presentaban menos errores en estas estructuras que los demás. Sin embargo, al cuantificar también los usos correctos de la estructura analizada, la autora comprobó que los hablantes de chino y japonés utilizaban la estrategia conocida como evitación y por ello habían producido muy pocas oraciones de relativo en total. De ahí que ese grupo de estudiantes presentaran menos errores en la estructura estudiada. Cabe señalar que, según Gass y Selinker (2008), el origen del fenómeno de la evitación todavía no está claro. Los autores indican que por un lado hay evidencia significativa de que el origen está en las diferencias entre la L1 y la L2, pero por otro, también hay evidencia de lo contrario, porque cuando hay muchas semejanzas entre las dos lenguas, los alumnos pueden creer que esto no es posible y pasan a evitar la utilización de determinada palabra o estructura lingüística.

De todos modos, la constatación de que los alumnos utilizan la estrategia de la evitación cuando una estructura les parece difícil, ha señalado la ineficacia del AE ya que enfoca exclusivamente en los errores que producen los estudiantes. Sin embargo, la identificación de la evitación no es tarea fácil y sólo se podría hablar de ello si se puede demostrar que el estudiante conoce la forma que está evitando y si los hablantes nativos de la L2 utilizan dicha forma en el contexto especificado (Ellis, 2008).

Otro problema encontrado por el AE ha sido identificar el origen del error. Según los presupuestos de este paradigma de estudio, los errores que podrían tener un origen en la L1 son llamados errores interlingüísticos, a diferencia de los errores intralingüísticos que podrían tener un origen en el proceso de aprendizaje de la lengua que se está estudiando (L2) como se ha dicho anteriormente. Sin embargo, no siempre es posible determinar con seguridad cuál es el origen del error. Así, Dulay y Burt (1974) establecieron una nueva categoría de errores denominada 'ambiguos', en la que figuraban los errores que no tenían un origen claramente identificable y que tenían características tanto de la L1 como de la L2.

Con todo, el Análisis de Errores tuvo su importancia y contribuyó a la investigación en ASL al mostrar, por ejemplo, que los errores que los estudiantes producen no son todos debido a la interferencia de la lengua nativa (L1). Además, cambió la visión negativa del error que era

visto antes como algo que se debía evitar y pasó a ser considerado parte del proceso de enseñanza-aprendizaje de una segunda lengua (Ellis, 2008).

En el campo de la enseñanza de español como lengua extranjera (ELE), destacan las investigaciones realizadas por a) Vázquez (1991), que ha analizado los errores orales y escritos que hacen parte del aprendizaje de español (L3) por estudiantes universitarios hablantes de alemán (L1), b) Santos Gargallo (1993), que ha examinado los errores en la expresión escrita en español (L2) de estudiantes universitarios serbocroatas y c) Fernández (1997), que ha estudiado los errores en la expresión escrita en español (L2) de hablantes de L1 variada (alemanes, japoneses, árabes y franceses) en una escuela oficial de idiomas en España.

Estos trabajos siguen siendo referencia para los investigadores que se dedican al estudio de los errores a través de la metodología del AE, como se puede comprobar en las memorias de máster de Rodríguez Paniagua (2001) sobre el aprendizaje de ELE por hablantes de árabe; Skjær (2004) sobre la comunicabilidad de los errores de ELE en lengua escrita por hablantes de inglés; Sánchez Jiménez (2006) sobre el análisis de los errores ortográficos de estudiantes filipinos de ELE; Mohd Hayas (2006) sobre los errores léxicos de estudiantes malayos de ELE, y en las tesis de Torijano Pérez (2002) sobre los errores gramaticales en el aprendizaje de ELE por lusohablantes; Sánchez Iglesias (2003) sobre el aprendizaje de ELE por hablantes de italiano; Fernández Jodar (2006) sobre los errores léxicos, morfosintácticos y gráficos en el aprendizaje de ELE por polacos, y la de Palapanidi (2012) sobre los errores léxicos de la lengua escrita en el aprendizaje de ELE por estudiantes griegos. Además, aunque se centre en el estudio de los errores léxicos en el ámbito de la enseñanza del inglés como lengua extranjera por alumnos hispanohablantes, destaca también como ejemplo de investigación actual que emplea la metodología de Análisis de Errores, el trabajo realizado por Agustín Llach (2011a) que investiga la relación entre los errores léxicos y la calidad de los textos.

Para James (1998), como fue imposible negar la importancia de los efectos del papel de la L1 en el aprendizaje de una L2, el Análisis de Errores no tuvo éxito y por esto dio lugar a la teoría de la Interlengua (IL), pero lo cierto es que el AE aporta una metodología que ayuda a tratar los errores de los alumnos y a la que los profesores de lenguas extranjeras o segundas



lenguas suelen recurrir con frecuencia en la búsqueda de una mejor comprensión del proceso de enseñanza-aprendizaje.

### 1.3. La Interlengua

Se suele considerar la Interlengua (IL) como la etapa que sucedió el Análisis Contrastivo (AC) y el Análisis de Errores (AE), pero a diferencia del AE que centra la atención solamente en los errores producidos por los estudiantes, la teoría de la IL toma en consideración el estudio de la producción del alumno como un todo, es decir, tanto los errores como lo que hace correctamente. Selinker (1969, 1972) ha sido el gran responsable del desarrollo de este concepto y de su difusión. Según este autor, la interlengua (IL) representa una fase más en el intento de aproximación al estudio global de la lengua producida por el estudiante durante el proceso de adquisición de una segunda lengua (L2).

El término interlengua se ha consolidado para referirse al sistema lingüístico en desarrollo del estudiante de una L2, pero antes de Selinker (1972), este concepto ya había sido tratado por otros investigadores como Corder (1967), que llamó a este sistema lingüístico de 'competencia transitoria', mientras Nemser (1971) lo hizo a través del término 'sistema aproximado'. La diferencia básica entre estas propuestas es la referencia que hace Nemser a la producción en L2 como 'sistema lingüístico desviado' (Santos Gargallo, 1993), denominación que no comparten los demás estudiosos, porque tanto para Corder (*op. cit.*) como para Selinker (*op. cit.*), la lengua del estudiante es un sistema lingüístico autónomo, que tiene su propia gramática, y es, además, un sistema correcto en sí mismo, que está en cambio continuo, es socialmente no reconocido y que puede ser descrito a través de un conjunto de reglas que es un subconjunto de las reglas de la gramática de la lengua meta (L2), cuyas oraciones no son desviadas ni erróneas, simplemente idiosincrásicas, y es además un sistema peculiar de un estudiante sólo o de un grupo de estudiantes con la misma formación académica y misma lengua materna (L1).

Una definición de interlengua bastante difundida y aceptada es la que proponen Frauenfelder *et al.* (1980: 46), que explica el sistema de la interlengua como

*un sistema lingüístico interiorizado que evoluciona tornándose cada vez más complejo, y sobre el cual el aprendiz posee intuiciones. Ese sistema es diferente del de la lengua materna y del de la lengua meta, y contiene reglas propias: cada aprendiz posee, en un estadio dado de su aprendizaje, un sistema específico<sup>5</sup>.*

Para Selinker (1972), las características propias de la interlengua se deben a diferentes razones y componen de forma individual el sistema lingüístico de cada estudiante. Esa lengua, a pesar de ser diferente en cada alumno, presentará características de intersección en aprendices con una preparación académica similar e idéntica L1. Las desviaciones se deben a las hipótesis creadas sobre el sistema de reglas y el funcionamiento de la lengua que se está aprendiendo. Después de la superación de cada etapa, hay errores que se corrigen y se superan, pero hay otros que revelan una tendencia a fosilizarse y que están presentes en la lengua de los estudiantes aún en los niveles avanzados de competencia. La fosilización (Selinker, 1978) o la estabilización (Han y Selinker, 2005), tema que se desarrolla con más atención en el capítulo 4, se da cuando el proceso de desarrollo de la interlengua se interrumpe, lo que se verifica cuando se detectan errores que persisten en la producción lingüística del estudiante de L2.

La transferencia lingüística, la hipergeneralización<sup>6</sup> (*overgeneralization*), la transferencia de instrucción, las estrategias de comunicación en la L2 y las estrategias de aprendizaje en la L2 son los procesos que constituyen el desarrollo de la interlengua, es decir, son las maneras a través de las cuales el estudiante internaliza el sistema de la L2 (Selinker, 1972), y se deben añadir a ellos los conceptos de comprobación de hipótesis (*hypothesis-testing*) y de fosilización, es decir, la persistencia de un determinado error en los diferentes estadios de desarrollo de la interlengua. Como se verá más adelante, en el capítulo 3 que trata del fenómeno del error, Corder (1967) sugiere que la comprobación de hipótesis es un proceso interno del estudiante que se hace evidente a través de los errores y que el cometer errores es una estrategia que utilizan tanto los hablantes de L1 como los de L2 para compensar las deficiencias del estadio concreto que atraviesa su interlengua.

Según Ortega (2009), la interlengua contiene tanto elementos de la L1 como de la L2, además de elementos que van más allá de estos dos como, por ejemplo, las ‘formas nuevas’

---

<sup>5</sup> Traducción de Fernández (1997).

<sup>6</sup> La “hipergeneralización” o “sobregeneralización” es un concepto que consiste en la generalización o extensión de una regla lingüística a ámbitos no apropiados (VanPatten y Benati, 2010).

(*new forms*) (Gass y Selinker, 2008). Así, la autora comenta ejemplos de interlengua en los que no se puede atribuir su causa ni a la L1 del estudiante ni al *input* que ofrece el entorno o los manuales, y explica que algunas soluciones que adoptan los estudiantes de L2 se observan también en la producción de niños que están aprendiendo su L1 y que no tienen, por lo tanto, conocimiento de otra lengua en la que apoyarse. Para la autora las formas de la interlengua son innovaciones sistemáticas provisionales y siendo así, la interlengua es mucho más que la suma del *input* de la lengua meta y la L1 del estudiante.

Dado que la interlengua es un sistema lingüístico en sí mismo que contiene elementos tanto de la L1 como de la L2 (Torras, 1994), un determinado aspecto de una interlengua sólo es un error si se compara con la L2 correspondiente. Así, James (1994) explica que, por un lado, algunas características de la IL se pueden explicar a partir de la comparación de los sistemas de la IL con la L2, como propone el AE, pero otras sólo se pueden explicar a través de la comparación de la IL con la L1, lo que el autor llama de análisis de la transferencia (AT). James (1998) también explica que los teóricos de la ASL suelen estudiar la interlengua *sui generis*, es decir, analizan el propio sistema lingüístico del alumno sin compararlo con la L1 de los estudiantes ni tampoco con la L2.

Tras analizar varias investigaciones que focalizan en el estudio del error en el ámbito de la enseñanza ELE, Alba Quiñones (2009) observa que investigadores como Santos Gargallo (1993) y Torijano Pérez (2002) se refieren al Análisis Contrastivo, Análisis de Errores y a la Interlengua como una sucesión continuada de teorías. Más concretamente, Torijano Pérez (2008) considera que las aportaciones de los diferentes modelos de estudio del error, el Análisis Contrastivo (AC), el Análisis de Errores (AE) y la Interlengua (IL), son complementarias entre sí, y constituyen la base del análisis y estudio del sistema lingüístico del estudiante. Sin embargo, Alba Quiñones (*op. cit.*) se inclina por el enfoque que adopta Sánchez Iglesias (2003), que defiende que esa idea provoca una confusión entre una hipótesis teórica (AC), un tipo de análisis de datos (AE) y una hipótesis psicolingüística (IL) y afirma que la interlengua es el objeto de interés de diferentes análisis, pero “no una tipología de análisis en sí misma” (Alba Quiñones, *op. cit.*: 2).

De cualquier modo, el análisis de la interlengua (IL), es decir, de la producción lingüística del estudiante de L2, viene generando contribuciones a la didáctica de la enseñanza de lenguas,

como bien apunta Carcedo González (1998), que afirma que desde perspectivas lingüísticas diversas, el análisis de la IL ha originado muchos trabajos empíricos que han focalizado en aspectos muy concretos del lenguaje y que ayudan a conocer los rasgos sistemáticos de cada fase del proceso de aprendizaje de una segunda lengua.

## 2. EL APRENDIZAJE DE UNA SEGUNDA LENGUA Y EL CONOCIMIENTO LINGÜÍSTICO

### PREVIO

Durante muchos años se ha hablado de la influencia de la lengua materna (L1) en el proceso de aprendizaje de una segunda lengua (L2). A lo largo del tiempo, este fenómeno ha recibido diversos nombres como transferencia, interferencia y también influencia croslingüística. En este trabajo se utilizan los términos transferencia lingüística, interferencia e influencia croslingüística para hacer referencia a los elementos de las lenguas previas que el estudiante ya conoce y que se observan en la producción lingüística de la lengua que está aprendiendo en un contexto formal de enseñanza-aprendizaje

Según Odlin (2003), la transferencia lingüística o la interferencia, se da tanto en contextos naturales de aprendizaje como en contextos formales. En este capítulo se comentan algunos de los trabajos relacionados con el fenómeno de la transferencia, se desarrollan estos conceptos, y por último se explica qué ocurre en el ámbito del léxico y las implicaciones de la distancia lingüística en los fenómenos de transferencia, especialmente en el caso de lenguas próximas como son el español (L2) y el portugués (L1), ya que es el contexto en el que se ha desarrollado este estudio.

#### 2.1. El estudio de la transferencia lingüística

La transferencia lingüística puede considerarse como “el impacto de un sistema de reglas sobre otro, del cual resulta la interlengua” (Vázquez, 1991: 44). En el presente apartado examinamos la transferencia lingüística desde su surgimiento con la hipótesis del Análisis Contrastivo (AC) bajo la perspectiva conductista, hasta la percepción de la transferencia lingüística como estrategia cognitiva, es decir, una estrategia de compensación en la que el estudiante recurre a un ítem de su L1 para mantener la comunicación por considerar que su interlocutor es capaz de comprenderlo. Se observa, por lo tanto, una gran variedad de matices en la trayectoria del concepto de transferencia, que van desde considerarla como factor esencial del proceso de construcción de la interlengua del estudiante de segundas lenguas (L2) (Lado, 1957), hasta la idea de que la transferencia no tenía mucha participación

en el desarrollo de la interlengua (Dulay y Burt, 1974). A continuación, hacemos un repaso de las principales líneas de pensamiento que han influido en este concepto.

### 2.1.1. Breve contextualización histórica

En los años cincuenta, las teorías de enseñanza-aprendizaje de segundas lenguas consideraban la transferencia lingüística como un factor de gran interés, porque se creía que la transferencia obstaculizaba el aprendizaje de una L2. Como se ha explicado anteriormente en el capítulo 1.1. (El Análisis Contrastivo), esta idea tiene su origen en la hipótesis del Análisis Contrastivo, que se fundamentaba en el conductismo y consideraba que el impedimento principal para la adquisición de la L2 estaba en los factores de interferencia creados por el conocimiento previo que tenía el estudiante del sistema de la lengua materna (L1).

La transferencia tuvo un papel determinante en la hipótesis del AC (Lado, 1957). El aprendizaje se consideraba bajo la perspectiva conductista, como un proceso de formación de hábitos y los errores demostraban la incapacidad para la creación de nuevos hábitos. Según Arabski (2006), los psicólogos conductistas fueron los que primero se refirieron a la transferencia técnicamente, hablando de los procesos automáticos, no controlados y subconscientes de los comportamientos del estudiante en el intento de producir las formas lingüísticas adecuadas a la lengua que está estudiando.

Después, los estudios de interlengua mostraron que no todos los errores producidos eran el resultado de interferencia lingüística, porque para que se pueda hablar con seguridad de interferencia, la forma producida en L2 debe contener rasgos de la L1 o de una L2 estudiada previamente. Además, se verificó que factores como la hipercorrección, la aplicación incompleta de las reglas de la L2 o la hipergeneralización (*overgeneralization*), que es la extensión de una regla lingüística de la L2 que se está estudiando a ámbitos no apropiados (VanPatten y Benati, 2010), explicaban también una parte considerable de los errores.

Según las teorías conductistas, el mayor obstáculo en el proceso de enseñanza-aprendizaje de una L2 era la interferencia del conocimiento previo. Como estas teorías hablaban que el aprendizaje era nada más que formación de hábitos, se creía que era necesario desaprender las antiguas costumbres para aprender las nuevas. Sin embargo, Ellis (2008) las critica

porque al aplicarlas a la enseñanza de L2, la idea de desaprender no tenía mucho sentido porque los estudiantes de una L2 no tenían que olvidar o desaprender su L1 para aprender una L2.

Se suponía que a mayor diferencia estructural entre las lenguas (L1 y L2), mayor sería la dificultad en la adquisición, puesto que la posibilidad de transferencia positiva es menor (Fernández González, 1995). El grado de dificultad, que se traducía en la cantidad de esfuerzo empleado para aprender un modelo en L2, estaba asociado a la similitud o diferencia entre la forma de la L1 y la de la L2 que se estaba estudiando (Ellis, 2008). Así, cuando las dos formas eran idénticas ( $L1=L2$ ), se consideraba que el aprendizaje se daba fácilmente, mientras que si las dos formas fueran diferentes ( $L1\neq L2$ ), aparecían dificultades y errores. De todas maneras, Vázquez (1991) resalta que relacionar similitud con facilidad y diferencia con dificultad significa adoptar una actitud simplista frente al complejo proceso de aprendizaje. Para la autora, la interferencia es, además de un proceso lingüístico, un fenómeno psicológico relacionado con la memoria.

En oposición a las teorías conductistas, y dentro del paradigma cognitivo de finales de los setenta, la transferencia se caracterizó como una estrategia que utilizaban los alumnos en el intento de solucionar problemas y comunicarse, utilizando el conocimiento de la L1 para solventar un problema de aprendizaje y/o comunicación en L2 (Kellerman, 1977; Sharwood Smith, 1979). De esa manera, esos autores defendían la idea de que el aprendizaje de lenguas debería pasar a ser considerado como un proceso creativo del cual se originaban las formas en la interferencia.

En el ámbito de la lingüística aplicada, 'transferencia' es un término utilizado para referirse a una estrategia de aprendizaje de una lengua extranjera en la que los estudiantes utilizan los conocimientos lingüísticos previos de su L1 y los aplican a la lengua que están aprendiendo. Esa estrategia puede ser una ventaja, si las dos lenguas tienen aspectos correspondientes, ya que así existe la posibilidad de que ocurra una transferencia directa de los elementos de la L1 a la interlengua del estudiante, ocasionando lo que se suele llamar 'transferencia positiva'. Sin embargo, también se dan casos de 'transferencia negativa' (o *interferencia*), cuando los patrones de las dos lenguas no coinciden (Arabski, 2006) y los estudiantes transfieren elementos o estructuras de la L1 que no corresponden a los de la L2.

También se verificó que “muchos de los errores cuyo origen se había buscado en la interferencia eran semejantes a los que cometen los niños al adquirir la lengua materna” (Galindo, 2004: 19), lo cual indicaba que no todos los errores se debían a la transferencia negativa de los elementos y estructuras de la L1, es decir, hay otros factores que influyen en el aprendizaje de segundas lenguas. Durante algún tiempo hubo un cierto rechazo de la influencia de la L1 en el aprendizaje de L2. Esto ocurrió principalmente porque había una necesidad de alejarse de las teorías conductistas del aprendizaje que habían considerado la interferencia un elemento clave para explicar la IL (Liceras, 1991).

Sin embargo, esa visión cambió y la atención de los investigadores volvió a centrarse en el papel de la lengua materna (L1) en el aprendizaje de L2. A partir de nuevos estudios, se verificó que el estudiante recurría tanto al conocimiento de su propia L1 como a sus conocimientos lingüísticos previos en el intento de facilitar el proceso de aprendizaje de la L2 (Ringbom, 1987). Los conocimientos previos se referían tanto a lo que el estudiante ya sabía sobre la lengua meta (L2), como aquello que sabía sobre su propia lengua materna (L1). Además, se verificó que los estudiantes de L2 recurrían no sólo a su L1, sino también a otras lenguas que hubiera estudiado previamente en el intento de facilitar el proceso de aprendizaje de la lengua meta (Sharwood Smith y Kellerman, 1986).

Odlin (1989: 27), a partir de las ideas de Kellerman y Sharwood Smith, define la transferencia lingüística como “la influencia resultante de similitudes y diferencias entre la lengua meta y cualquier otra lengua que haya sido adquirida anteriormente, aunque ese aprendizaje haya ocurrido de manera imperfecta”. El autor argumenta que la transferencia no es una consecuencia de formación de hábitos ni una simple aplicación de las características de la L1, sino que el proceso de transferencia de L1 a L2 es parte de un proceso general del que están implicados tanto los contenidos lingüísticos como el conocimiento de mundo, los aspectos pragmáticos, los conocimientos de otras lenguas extranjeras, etc.

Así, se consolidaba la idea de que el estudiante de L2 no empezaba su aprendizaje desde cero y que recurrir a su L1 (y a otros conocimientos lingüísticos previamente adquiridos) en la búsqueda de facilitar el proceso aprendizaje de una L2 era una estrategia muy utilizada por los estudiantes que pasó a ser valorada positivamente.



Como se ha visto, la transferencia puede ser considerada tanto un proceso de aprendizaje como un proceso de producción (Kohn, 1986). Como proceso de aprendizaje, la transferencia es parte del comportamiento interlingüístico del alumno que incluye la transformación creativa de la información lingüística del *input* (véase nota al pie 1, p. 11) en conocimiento interlingüístico. La transferencia como proceso de producción está involucrada en la producción de *output*<sup>7</sup> significativo a través de la recuperación del conocimiento y de los esfuerzos para superar las lagunas lingüísticas que no pueden esquivar a través de la estrategia de la evitación (Schachter, 1974), es decir, cuando el estudiante no utiliza determinado elemento o estructura lingüística en su producción en L2 porque no está seguro de ello.

### 2.1.2. Variedad terminológica

Tras ese repaso de la transformación que ha tenido el estudio de la transferencia lingüística desde su surgimiento en los años cincuenta, con el Análisis Contrastivo (AC) y apoyada por teorías conductistas, se han propuesto nuevas terminologías que incluyeran las nuevas perspectivas relacionadas con la idea de transferencia y consideraran no sólo los fenómenos de transferencia lingüística, sino también que se propusieran a analizar cómo el conocimiento lingüístico previo y la interlengua interactúan en el proceso de aprendizaje.

Corder (1983) había sugerido el empleo de términos alternativos como “influencia de la lengua materna” porque lo consideraba un término neutro y amplio para hacer referencia a lo que se había llamado comúnmente de ‘transferencia’. Para el autor, la teoría original de transferencia asignaba un papel muy restrictivo a la L1 que no cubría todo el fenómeno de forma suficiente. Así, bajo el término propuesto, fenómenos como la transferencia, la interferencia, el préstamo y la evitación estarían incluidos.

Otros autores como Gass y Selinker (1983) y Odlin (1989) se referían al fenómeno de la transferencia lingüística a través del término ‘transferencia de lengua’ (TL) (*language transfer*). Sin embargo, para Odlin (*op. cit.*), la transferencia se refiere específicamente a la influencia que resulta de las semejanzas y diferencias entre la lengua meta y otras lenguas

---

<sup>7</sup> *Output* es un término que hace referencia a la lengua que produce el estudiante (Centro Virtual Cervantes, 2012).

que el estudiante haya adquirido anteriormente, y puede ocurrir en los niveles estratégico, discursivo, lingüístico y pragmático.

También con el objetivo de incluir fenómenos como la transferencia, la interferencia, la evitación, el préstamo y aspectos relacionados a los efectos de la L2 en la L1, y analizar cómo el conocimiento lingüístico previo y la interlengua interactúan en el proceso de aprendizaje, Kellerman y Sharwood Smith (1986) propusieron el término 'influencia croslingüística' (ICL) (*crosslinguistic influence*).

Sin embargo, James (1994) afirma que el término ICL sólo pretendía diferenciar el producto, la transferencia, del proceso, la influencia. Así, el autor utiliza el término 'análisis de la transferencia' (AT) (*transfer analysis*) al referirse a la comparación entre los sistemas lingüísticos de la interlengua (IL) y de la L1 y argumenta que el AT difiere tanto del Análisis Contrastivo (AC) como del Análisis de Errores (AE), una vez que el AT no compara L1 con L2 como en el AC, ni tampoco se está comparando IL con L2 como en el AE. Para el autor, se pueden explicar algunas características de la IL a partir tanto de su comparación con la L2, como con la L1. En la figura 2 vemos un resumen de los sucesivos paradigmas de aprendizaje de L2.

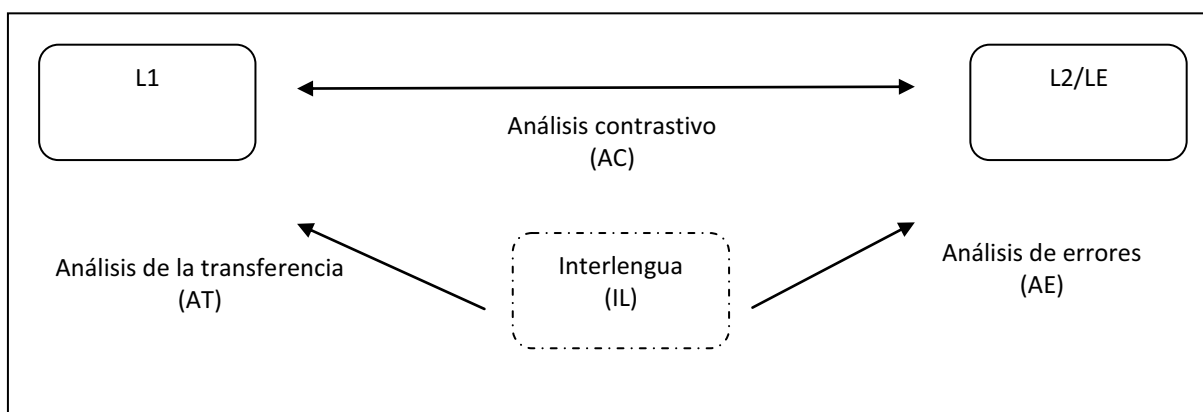


Figura 2. Los paradigmas de aprendizaje de L2/LE.

Kellerman y Sharwood Smith (*op. cit.*) defienden la idea de que el término influencia croslingüística (ICL) pueda usarse para referirse al fenómeno de la transferencia sin importar la dirección de la influencia ( $L1 \leftrightarrow L2$ ). Así, se estudian no sólo los fenómenos de influencia de la L1 en el proceso de aprendizaje de una L2, sino también la influencia de la L2 en la lengua materna del estudiante. Además, el término acoge los estudios tanto en adquisición de segundas lenguas como los de lengua extranjera, e incluye los variados contextos de

aprendizaje como el natural y formal. Esos autores señalan que la ICL constituye un aspecto particular de investigación en ASL que procura estudiar cómo la lengua materna (L1) y las segundas lenguas (L2) interactúan en el proceso de adquisición.

En esa misma línea, Ellis (2008) afirma que ninguna teoría del aprendizaje puede descartar la acción de la transferencia de formas de la L1 a la L2; pero destaca que del antiguo concepto conductista de transferencia como interferencia negativa de hábitos adquiridos en la L1 sobre el aprendizaje de una L2, se ha pasado a una concepción más amplia de influencia interlingüística, que comprende tanto aspectos negativos (interferencias) como positivos. Además, con la utilización del término ICL se hace hincapié no sólo en la vertiente pasiva de la transferencia (visible sobre todo en los sectores del lenguaje menos sometidos a control consciente como la fonética), sino también en los activos o estratégicos, es decir, la posibilidad de que el estudiante utilice la L1 como fuente de hipótesis sobre la L2, con finalidad tanto comunicativa como de aprendizaje.

Se han realizado estudios sobre la influencia croslingüística en todos los niveles lingüísticos (fonológico, léxico, semántico y sintáctico) como se puede ver en la revisión bibliográfica que hace Liu (2001). La ICL ha sido relacionada con la predominancia lingüística<sup>8</sup> (*language dominance*), la psicotipología<sup>9</sup> (*psychotypology*), la edad, el grado, la competencia estratégica y el nivel de competencia en la lengua meta y actualmente ha recibido atención creciente principalmente en las investigaciones que se realizan en contextos multilingües (Cenoz *et al.*, 2003).

### 2.1.3. Niveles lingüísticos en los que se verifica la transferencia

La transferencia es un fenómeno que se manifiesta en diferentes niveles lingüísticos: fonético, fonológico y léxico. Sin embargo, todavía es un fenómeno incierto en los niveles gramaticales y sintácticos, aunque hay evidencia del fenómeno en la adquisición del orden

---

<sup>8</sup> La predominancia lingüística se refiere al hecho de que en un individuo que conoce más de una lengua, una de ellas prevalece sobre la otra.

<sup>9</sup> La psicotipología es un concepto difundido por Kellerman (1977, 1979, 1983) que se refiere a la percepción que tiene el estudiante de la distancia lingüística entre la L1 y la L2.

de palabras y en los sintagmas verbales. Asimismo, se ha verificado evidencia de la transferencia de parámetros textuales y pragmáticos de la L1 a la L2 (Bou, 1998).

Además de los niveles lingüísticos ya referidos, la transferencia también es muy productiva en niveles cognitivos de referencia, como el conocimiento del mundo, la comunicación no-lingüística y en las habilidades propias del aprendizaje como hacer asociaciones, comparaciones y formulación de hipótesis. Esas variables lingüísticas y cognitivas se combinan de forma diferente en cada momento del proceso de enseñanza-aprendizaje y estimulan la posible ocurrencia de la transferencia (Ainciburu, 2007).

Odlin (2006) pone atención a fenómenos de restricción en la transferencia, como puede ser el caso de los morfemas ligados (sufijos, prefijos, morfemas verbales), el orden básico de palabras, la proyección funcional y los modismos. Según Dusková (1984), la restricción en la morfología de ligación<sup>10</sup> (*bound morphology*) depende mucho más de las lenguas en contacto y de las percepciones de los estudiantes sobre la distancia croslingüística.

#### 2.1.4. Condiciones que favorecen la transferencia

De acuerdo con Vázquez (1991: 45), para que se produzca interferencia son necesarias cuatro condiciones indispensables: a) que haya convergencia de dos o más lenguas cuyas reglas no sean del todo equivalentes; b) un hablante que crea que es aceptable transferir de una lengua a otra; c) un contexto en el que la interferencia sea posible, y d) una percepción de las estructuras en términos de similitud y contraste.

Se han verificado regularidades que gobiernan la transferencia. Ésta depende de características personales del sujeto como edad, motivación, estado de desarrollo de su interlengua o percepción de la distancia entre L1 y L2. Además, la intensidad de la transferencia lingüística depende también del nivel de competencia del estudiante (Arabski, 2006).

Jarvis (2000: 26) defiende la idea de que sólo controlando las variables se podrá descubrir cuándo la transferencia ocurrirá y bajo qué condiciones. Para que eso sea posible, este

---

<sup>10</sup> Los morfemas de ligación son aquellos que sólo aparecen junto a otros morfemas, son dependientes, es decir, se unen a otras palabras para añadir información a su significado (ej.: prefijos, sufijos, infijos).

autor propone una lista de elementos adaptada de la propuesta de Ellis (2008) que se deben tener en cuenta al realizar estudios que se dediquen a investigar la transferencia y que permite controlar las variables:

- a) edad
- b) personalidad, motivación y aptitud lingüística
- c) educación social y bagaje cultural
- d) bagaje lingüístico (L1s y L2s previas)
- e) tipo y cantidad de exposición a la lengua meta
- f) nivel de conocimiento en la lengua meta
- g) distancia lingüística entre L1 y la lengua meta
- h) tipo de tarea y área de uso de la lengua
- i) prototipicalidad<sup>11</sup> (*prototypicality*) y diferenciación<sup>12</sup> (*markedness*) del aspecto lingüístico.

Los tres primeros ítems se refieren a factores extralingüísticos. El interés en controlar la edad de los participantes se debe a que el comportamiento de los niños difiere del de los adultos. Otro factor que constituye un objeto de estudio en sí mismo es la motivación y los factores que la condicionan. Conocer la motivación y la aptitud lingüística de los participantes ayuda en la interpretación de los resultados. Los demás ítems tienen relación directa o indirecta con factores lingüísticos. El conocimiento lingüístico previo de los estudiantes influye en el aprendizaje de una nueva L2 (Ringbom, 1987). La calidad y la cantidad de la exposición a la lengua meta, conocido como el *input* lingüístico comprensible (Krashen, 1981) es determinante en el aprendizaje de L2. La importancia de controlar el nivel de conocimiento de la L2 que tiene el participante se debe a que la transferencia incide de diferentes maneras en la interlengua a lo largo del proceso de aprendizaje de una L2 (ej.: Celaya y Torras, 2001; Navés *et al.*, 2005; Celaya, 2007).

---

<sup>11</sup> Prototipicalidad es un concepto que tiene su origen en la psicología, destacablemente con Eleanor Rosch. En la lingüística cognitiva se refiere a la idea que determinado ítem está más próximo a la definición de la categoría que otros. En la lingüística de corpus, generalmente se refiere a la frecuencia de determinado ítem en la muestra (Gilquin, 2006).

<sup>12</sup> Diferenciación o *markedness* es un concepto lingüístico que se refiere a la presencia de un determinado fenómeno o forma lingüística. Si, por un lado, el fenómeno es muy común, se dice que es una característica no marcada o poco marcada (*unmarked*). Por otro lado, si la característica no es muy frecuente, se dice que es marcada (*marked*) (VanPatten y Benati, 2010).

En lo que se refiere a la distancia de las lenguas implicadas en el aprendizaje, como se verá más adelante en el apartado 2.3. (La transferencia y la proximidad lingüística), los estudiantes tratan de buscar semejanzas entre su propia lengua (L1) y la lengua que están estudiando (L2) (Ringbom y Jarvis, 2009). Es interesante destacar que las estructuras marcadas de la L1 son menos propensas a la transferencia. Sin embargo, el caso de las expresiones idiomáticas, estructuras no marcadas, es diferente quizás por su carácter metafórico, como apunta Kellerman (2000). Asimismo, Laufer (2000) verificó que el nivel de competencia en L2 y la proximidad entre la L1 y la L2 son factores que influyen en la transferencia de las expresiones idiomáticas.

El contexto de uso de la lengua y el medio, oral o escrito, implicado en la tarea también juegan un papel importante en las estrategias que utilizan el estudiante y consecuentemente en el tipo de transferencia que se observa luego. Para finalizar, se resalta que hay tres factores que determinan la transferencia: a) la percepción de la especificidad de un determinado ítem en la L1 (prototipicalidad) (Kellerman, 1978, 1986); b) la percepción de la distancia tipológica entre la L1 y la L2 (psicotipología), y c) el nivel de competencia que tiene el estudiante en la L2.

Para seguir profundizando en el estudio de la transferencia lingüística, dado que nuestro objetivo apunta al léxico, a continuación se comentan los estudios que se han dedicado a conocer mejor los fenómenos de transferencia en el ámbito del léxico.

## 2.2. La transferencia lingüística y el léxico

Se considera el léxico como uno de los aspectos lingüísticos más permeables de la L1 y la transferencia lingüística es una estrategia compensatoria muy utilizada por los estudiantes de segundas lenguas (L2) (Agustín Llach, 2007) que les ayuda a mantener la comunicación con su interlocutor.

De acuerdo con el MCER (2002: 5.2.1.1.), la competencia léxica es “el conocimiento del vocabulario de una lengua y la capacidad para utilizarlo, se compone de elementos léxicos y elementos gramaticales”. De acuerdo con Leffa (2000), para conocer el léxico de una lengua es necesario disponer de recursos cognitivos, afectivos y especialmente de tiempo. El autor

afirma que cuando se habla de léxico, generalmente se hace referencia a la totalidad de las palabras de una lengua, es decir, al saber interiorizado por parte de los hablantes de una comunidad lingüística, y que vocabulario es una parte del léxico que representa una determinada área de conocimiento.

Los estudiantes consideran el vocabulario como un elemento fundamental en el aprendizaje de lenguas y muchas veces equiparan la lengua al vocabulario (James, 1998). Investigadores como Saville-Troike (1984: 199) confirman esta percepción de los estudiantes y afirman que el conocimiento del vocabulario es el área más importante de la competencia en L2 con relación al rendimiento académico. Asimismo, para los estudiantes, disponer de un vocabulario amplio es fundamental para escribir textos académicos, dado que la ampliación del vocabulario es un elemento determinante en el desarrollo de la expresión escrita (Leki y Carson, 1994). Según Leffa (2000), lo que existe entre el texto y la palabra es un proceso de interacción que se basa en reglas fundamentales, en las que el texto proporciona el contexto en el que uno de los significados semánticos de la palabra prevalece sobre los demás.

Como el léxico es el conjunto de palabras de una lengua, es necesario definir qué es una palabra. Según el *Gran Diccionario de Uso del Español Actual* (Sánchez Pérez, 2001), una palabra es una “agrupación de sonidos o letras que constituye una unidad con significado y función propias en un sistema lingüístico”. Asimismo, según el *Diccionario de la Lengua Española* de la Real Academia Española (DRAE, 2001) una palabra es un “segmento del discurso unificado habitualmente por el acento, el significado y pausas potenciales inicial y final”. Para Gómez Molina (1997), aunque la unidad léxica por excelencia es la palabra, existen otros tipos de unidades léxicas (lexías complejas, locuciones idiomáticas, etcétera) o expresiones fijas, que el estudiante de ELE ha de ser capaz de utilizar cuando el acto comunicativo lo requiera.

Sin embargo, Goes de Andrade (2006: 1) explica que

*saber uma palavra implica conhecer sua forma e seu significado, assim como seus possíveis sentidos a partir do estabelecimento de uma rede de relações com outras palavras, em virtude da experiência cotidiana com a língua em seus diferentes registros e canais. Além disso, saber uma palavra implica também conhecer as restrições no seu uso, do ponto de vista morfo-sintático, léxico-semântico e pragmático-cultural.*

Tras conocer qué es una palabra y qué implica saber una palabra, vemos que las palabras se suelen agrupar en dos categorías: léxica y funcional. Por un lado, se conocen como palabras léxicas aquéllas que tienen significado, contenido semántico, y comprenden los sustantivos, adjetivos, verbos y adverbios. Estas palabras constituyen inventarios abiertos, es decir, como una lengua es un sistema vivo y dinámico, siempre surgen nuevos términos que se van incorporando. Por lo tanto, el léxico se caracteriza por ser un sistema abierto, irregular y no sistemático. Por otro lado, las palabras funcionales son aquéllas que tienen un significado gramatical, es decir, establecen las relaciones entre las palabras léxicas e incluyen los artículos, pronombres, preposiciones, conjunciones y verbos copulativos. Las palabras funcionales, al contrario de las léxicas, constituyen inventarios cerrados, porque sólo hay un número finito de cada una de ellas (Sánchez, 2002).

En lo que se refiere a la transferencia léxica, según Kellerman (1995), se ha identificado evidencia de la influencia del léxico de la L1 en la interlengua de los estudiantes de L2, por un lado porque ellos utilizan la transferencia léxica como estrategia de compensación, como se ha explicado anteriormente, y por otro porque el léxico es uno de los aspectos lingüísticos más permeables.

Ringbom (1986) es uno de los investigadores que más ha contribuido en el estudio de la transferencia léxica y la define como la aplicación de hipótesis de aprendizaje que consideran los ítems lexicales como equivalentes traducidos, o que tienen las mismas características semánticas, de los ítems de la L1 del estudiante u otra lengua que conozca bien. En su estudio comparativo entre grupos de estudiantes de inglés como lengua extranjera (ILE) con alumnos de distintas L1 (finlandés y sueco), Ringbom (1987) encontró una gran predominancia de la influencia de la L1 en el léxico. El autor argumenta que la semejanza croslingüística entre la L1 y la L2 puede ser asumida como un elemento importante en el almacenamiento de ítems léxicos y que este conocimiento es como un sistema o un conjunto de sistemas que se puede utilizar tanto para la comprensión como para la producción lingüística. También explica que además de la similitud entre ítems individuales hay que considerar otros factores que interfieren en la transferencia léxica.

Aparte de los estudios de Ringbom, otros investigadores han colaborado para una mejor comprensión de los fenómenos de transferencia léxica. Los estudios centrados en la



competencia concluyen que estudiantes con baja competencia tienden a transferir más elementos de su L1 que los estudiantes más avanzados (Celaya y Torras, 2001; Woodall, 2002; Navés *et al.*, 2005; Celaya, 2007). Sin embargo, otros estudios demuestran que la influencia de la L1 aumenta a lo largo del desarrollo de la interlengua, es decir, que cuanto más competentes son los estudiantes, más transfieren, especialmente con respecto al número de préstamos usados (Sanz, 2000). Asimismo, hay estudios que indican que la influencia ni aumenta ni disminuye a lo largo del desarrollo de la interlengua (Poullisse, 1990), mientras otros afirman que este proceso fluctúa hasta tomar una dirección específica (Engber, 1995; Jarvis, 1998).

Con base en esos resultados tan contradictorios, Jarvis (2000) cuestiona si todos esos estudios están evaluando el mismo fenómeno y reivindica que se establezcan unos criterios metodológicos específicos para el estudio de la transferencia léxica. Su propuesta consta de tres criterios (Jarvis, *op. cit.*: 5): a) una definición imparcial de la influencia de la L1 (transferencia) que funcione como marco metodológico para estudios de este tipo; b) una declaración sucinta pero exhaustiva de los tipos de evidencia que deben considerarse cuando se presentan argumentos a favor o en contra de la influencia de la L1, y c) una lista de variables externas que se deben controlar en cualquier investigación rigurosa sobre la transferencia.

Investigaciones con estudiantes bilingües han analizado la influencia croslingüística (ICL) con relación a la predominancia de la lengua y el nivel de enseñanza en contextos escolares. El estudio realizado por Navés *et al.* (2005) ha investigado si la ICL, analizada a través de préstamos e invenciones léxicas, cambia con relación al grado de instrucción en lenguas relacionadas tipológicamente (cantidad de instrucción en L2) y cómo ésto influye en la transferencia de palabras contenido y palabras funcionales en estudiantes poco proficientes (L2 = inglés) hablantes bilingües de lenguas relacionadas tipológicamente (español y catalán). Las autoras concluyen que los estudiantes de niveles avanzados hacen menos uso de los préstamos y de las invenciones léxicas que los estudiantes en niveles más bajos. Ese resultado parece confirmar estudios previos sobre transferencia como estrategia compensatoria que afirmaban que los estudiantes menos proficientes presentaban más ICL. Sin embargo, en el estudio de Cenoz (2001) el resultado era el opuesto, pues sólo la mitad de los estudiantes utilizaba esta estrategia.

Los estudios de Jiang (2002) revelan que la transferencia léxica positiva por parte de estudiantes de L2 con la palabra cognada correspondiente en L1 se debe, no sólo a la similitud de las formas léxicas, sino también a la disponibilidad de equivalentes traducidos que permiten una localización fácil de las formas de L2 en lemas de la L1. Un factor clave en la transferencia léxica es poder averiguar si los estudiantes forman una nueva descripción semántica para una palabra de la L2 o si trazan un mapa en la representación semántica existente de su L1 (Ellis, 2008).

En el ámbito de la enseñanza de ELE, Vázquez (1991) afirma que, con referencia al léxico, los verbos que menos resistencia ofrecen a la interferencia son los reflexivos, especialmente aquéllos que presentan dos formas, una reflexiva y otra no, con los cambios semánticos correspondientes. Asimismo, las desviaciones semánticas y pragmáticas son las que tienen un índice de inteligibilidad más bajo. Por otro lado, las morfosintácticas, son las que tienen un índice más alto. A mayor grado de inteligibilidad en la L2, congruencia entre L1 y L2 y productividad en L2, menor grado de resistencia a la interferencia. La resistencia de una estructura a la interferencia está en estrecha relación con la 'fossilización' (véase capítulo 4), ya que las estructuras que presentan baja resistencia a la interferencia se fosilizan con mayor facilidad.

Los préstamos han sido analizados en muchos estudios con relación a palabras contenido y palabras funcionales. Uno de los estudios averiguó que estudiantes holandeses de inglés como lengua extranjera (ILE) de 9º y 11º grado transferían más palabras funcionales (es decir, palabras de categorías gramaticales cerradas como preposiciones, conjunciones, pronombres, etcétera...) que palabras contenido (palabras de las categorías gramaticales abiertas: sustantivo, adjetivo, verbo, adverbio) de su L1 en tareas orales (Poullisse y Bongaerts, 1994). Sin embargo, el estudio realizado por Cenoz (2001) demuestra que estudiantes bilingües vasco-español de inglés como L3 transferían más palabras contenido que palabras funcionales en tres niveles de competencia distintos (2º, 6º y 9º curso).

Celaya y Torras (2001) estudiaron el papel de la L1 (español y catalán) en palabras de clase abierta en ILE para analizar las diferencias en el número y en los *types* de transferencia léxica en sujetos agrupados en tres edades distintas: niños, adolescentes y adultos. Sus resultados sugieren que los niños recurren más a la L1 cuando la transferencia es

ortográfica, mientras que los adolescentes y adultos recurren más a la L1 que los niños en los procesos que combinan el conocimiento de L1 y L2 (acuñación). Sin embargo, destacan que la influencia de la L1 en el vocabulario escrito en ILE puede verse afectado no sólo por la edad, sino también por las diferentes metodologías que tienen lugar en las prácticas en clase.

En un estudio longitudinal, Celaya (2006) analiza la relación entre la influencia de dos L1 (español y catalán) y los niveles de competencia en L2 en palabras de clase abierta en el ámbito del inglés como lengua extranjera. Los resultados revelan que aunque la influencia de la L1 disminuye cuando los niveles de competencia en L2 aumentan, hay una forma no estándar (calcos) que no disminuye tras 726 horas de instrucción. Así, la autora sugiere que la influencia de la L1 y la competencia de la L2 están relacionadas de varias maneras dependiendo del tipo de la palabra no estándar.

En esos dos últimos estudios al que acabamos de referirnos, y que son parte del proyecto *Barcelona Age Factor* (BAF) en la Universidad de Barcelona, las investigadoras analizaron los datos con base en la clasificación de errores de James (1998), creando una taxonomía muy útil para clasificar la transferencia léxica en cuatro categorías: error ortográfico (*misspelling*), préstamo (*borrowing*), acuñación (*coinage*) y calco (*calque*).

De acuerdo con Ringbom (2006), en los estadios iniciales del proceso de enseñanza-aprendizaje, los estudiantes tienden a focalizar más en la forma que en conceptos más abstractos de significado o función, y las semejanzas formales percibidas ayudan a establecer las relaciones croslingüísticas en la memoria a largo plazo. El autor afirma también que se ha averiguado que la interferencia ortográfica interlingüística aumenta proporcionalmente dependiendo de la relación existente entre las dos lenguas (L1 y L2), es decir, cuánto más próximos sean los idiomas, se observarán más errores ortográficos con origen en la L1 del estudiante, pero resalta que se tratan generalmente de errores individuales.

### 2.3. La transferencia y la proximidad lingüística

Este trabajo se propone conocer el papel de la L1 en los procesos de transferencia, en el caso específico de dos lenguas que comparten una gran cantidad de rasgos lingüísticos y culturales como es el caso del portugués (L1) y del español (L2). Como se ha visto, la L1 tiene un papel importante en la transferencia, pero su influencia depende de muchos factores. Uno de ellos se debe a la distancia/proximidad lingüística entre las lenguas en contacto. Otro factor se refiere a la percepción de la distancia/proximidad de las lenguas por parte de los hablantes (psicotipología) (Kellerman, 1978), que se ha explicado anteriormente en el apartado 2.1.2. Variedad terminológica). Por ello, a continuación se explora lo que se sabe hasta el momento sobre la distancia lingüística y la percepción de ésta por parte de los estudiantes de segundas lenguas (L2).

En el apartado dedicado al Análisis Contrastivo, se ha visto que los estudiantes de una lengua extranjera recurren frecuentemente a su lengua materna (L1) durante el proceso de aprendizaje de una L2. Según Ringbom y Jarvis (2009), los alumnos recurren a la L1 en búsqueda de las semejanzas entre su propia lengua, las lenguas que conocen previamente y la lengua que están estudiando. Con todo, según esos autores, las investigaciones en ASL no han puesto tanta atención a las semejanzas entre las lenguas como han hecho con las diferencias existentes.

Según Braga (2010), las oportunidades de utilización de dicha estrategia serán mayores cuánto más próximas tipológicamente sean las lenguas en contacto, aunque eso no sea así necesariamente. Esto debido a que no es sólo la cercanía lingüística lo que está en juego, sino también la percepción de esta vecindad por parte del hablante. Más adelante, en el apartado siguiente, se profundizará en ese concepto.

Como se ha comentado anteriormente en el apartado 2.1. (El estudio de la transferencia lingüística) referente a la transferencia lingüística, se suele referir a los resultados del proceso de transferencia lingüística como transferencia positiva o negativa (interferencia). La idea era la de que los estudiantes transferían rasgos de su L1 a la lengua que estaban aprendiendo y se consideraba transferencia positiva cuando el alumno utilizaba estructuras de la L1 similares a las estructuras de la L2, y la transferencia negativa (interferencia) se daba cuando el alumno hacía uso de alguna estructura de la L1 que no correspondía a la

estructura de la L2. Sobre ese tema, Larrinaga (2006) afirma que los cognados y los falsos cognados son ejemplos claros de lo que pueden ser los mecanismos de transferencia positiva y negativa. Además, el autor explica que los falsos cognados pueden inducir al estudiante a una transferencia negativa, dado que la semejanza en la forma puede pasar la falsa idea de que los ítems corresponden también en el ámbito del significado (error interlingüístico). Sin embargo, para Gass y Selinker (2008) no existen dos procesos de transferencia, uno positivo y otro negativo, sino que la transferencia es un proceso único y el resultado (positivo o negativo) de la transferencia se verificaría una vez analizada la producción del estudiante, como se ha explicado en el apartado 1.1. (El Análisis Contrastivo).

Es importante resaltar que la distancia o la proximidad lingüística pueden ser consideradas tanto un fenómeno lingüístico como psicolingüístico (Ellis, 2008). Ringbom y Jarvis (2009) utilizan los términos 'semejanza real' (*actual similarity*), para referirse al primero, y 'semejanza asumida' (*assumed similarity*), para referirse al segundo. Como fenómeno lingüístico, la distancia se establece de acuerdo con el grado de diferencia entre las dos lenguas. Como fenómeno psicolingüístico, para establecer la distancia, hay que conocer qué percepción tienen los estudiantes del grado de diferencia entre su L1 y la lengua meta (L2).

Kellerman (1995), Jarvis (2000) y Ringbom (2006) investigaron la transferencia enfatizando tanto la diferencia como la semejanza lingüística. Aunque algunos estudios se hayan centrado en las diferencias entre las lenguas, se ha constatado que inconscientemente los estudiantes buscan las semejanzas en las lenguas previas que conocen y la que están aprendiendo. Asimismo, esos autores destacan la importancia de la semejanza croslingüística en el aprendizaje de una lengua extranjera y afirman que la L1 del estudiante puede facilitar el aprendizaje de la L2 cuando hay semejanzas entre la L1 y la L2. Además, añaden que el efecto positivo de la L1 es innegable cuando las dos lenguas en cuestión comparten un gran número de cognados.

La intensidad y el tipo de interferencia va a depender de qué lenguas están en contacto (L1 y L2), es decir, la transferencia lingüística se dará con más intensidad cuando los dos sistemas lingüísticos sean genéticamente próximos, lo que significa que habrá una mayor cantidad de elementos en común (Arabski, 2006). Desde el punto de vista semántico, las

palabras que tienen correspondencia casi automática con la L1 tienen baja resistencia a la transferencia (Vázquez, 1991). Un ejemplo de ello es el estudio realizado por Ringbom (1987) sobre la transferencia en el aprendizaje de inglés por hablantes de finlandés (L1) y por hablantes de sueco (L1), que compartían el mismo contexto cultural y educativo. Se ha verificado que los hablantes nativos de sueco tenían más ventajas durante el aprendizaje del inglés que los hablantes nativos de finlandés debido a la semejanza lingüística que existe entre el sueco y el inglés.

Cenoz (2001) también discutió un número de factores que pueden afectar a la influencia croslingüística en general como, por ejemplo, la edad, el contexto de uso y el nivel de competencia. La investigadora ofreció evidencia empírica de que la distancia lingüística es un factor relevante en la adquisición de inglés por bilingües españoles y vascos. Algunos de los participantes de su estudio tenían como lengua dominante el vasco y otros, el español. Con relación a la distancia lingüística, el idioma vasco no está relacionado con el español ni con el inglés. En los resultados de su estudio, la investigadora encontró más evidencia de transferencia del español al inglés que del vasco al inglés. Ese fue el caso para todos los estudiantes independientemente de la lengua dominante. Con relación al efecto de la edad, los niños mayores presentaron más influencia croslingüística que los niños más jóvenes. También hay factores relacionados con la lengua, con más transferencia de palabras contenido que palabras funcionales. Un resultado interesante fue el de que cuando las palabras en inglés son extranjerizadas, sólo las palabras en español estaban en la base y no las palabras vascas.

En el ámbito de la enseñanza de ELE a brasileños se ha identificado que en los niveles bajos de competencia, la proximidad lingüística es un elemento positivo porque permite que la comunicación se realice en la lengua meta desde el primer día de clase (Camorlinga, 1997; Espiga, 1999). Además, los alumnos se sienten motivados y participan activamente en el aula. Con todo, a partir del nivel intermedio, se nota un cambio en el progreso de la interlengua, y la facilidad y la rapidez en el aprendizaje empiezan a dar lugar no sólo a rasgos fosilizables sino también a la frustración (Andrade Neta, 2000; Calvi, 2004; Durão, 2004; Calvo Capilla, 2010). De todos modos, hacen falta más estudios que exploren en qué medida la proximidad lingüística entre la L1 y la L2 favorece el aprendizaje de la L2.

## 2.4. La percepción de la proximidad lingüística

Algunos investigadores señalan que el fenómeno de la transferencia está más relacionado con la percepción que tienen los estudiantes de segundas lenguas, que con la distancia lingüística existente entre las lenguas en contacto (Kellerman, 1978; Odlin, 1989, 2006). Esto se debe a que, aparte de los factores lingüísticos, los factores psicológicos inherentes a cada individuo pueden tanto generar la interferencia de la L1 como crear resistencia al aprendizaje de la L2 (Vázquez, 1991).

De acuerdo con los supuestos del Análisis Contrastivo, cuanto mayor sea la distancia entre las lenguas, más errores se van a producir. Por otro lado, numerosas investigaciones sobre las producciones de estudiantes (ej.: Schachter, 1974; Ringbom, 1987) demuestran que la interferencia se produce con más frecuencia entre aquellas lenguas y entre estructuras lingüísticas que el estudiante siente como más próximas. Además, la interferencia ocurre preferentemente cuando la lengua meta (L2) tiene una estructura parecida a la de la lengua materna (L1), lo que propicia la hipergeneralización (Fernández, 1997).

Ringbom es uno de los estudiosos que se han dedicado a investigar cómo la proximidad lingüística interfiere en el aprendizaje de una L2 y ha identificado una serie de características propias de este contexto. El autor establece una distinción que se basa en la percepción que el estudiante tiene de la similitud lingüística y, por un lado, define como ‘influencia croslingüística abierta’ (*overt cross-linguistic influence*) las semejanzas percibidas, o conscientes, y puede clasificarse en transferencia y préstamo, y por otro, llama de ‘influencia croslingüística encubierta’ (*covert cross-linguistic influence*) el fenómeno que ocurre cuando el estudiante no es consciente de la semejanza (Ringbom, 1986).

Cuando el estudiante percibe una semejanza entre las palabras y estructuras de las lenguas en contacto, se observa un efecto general facilitador en la comprensión y aprendizaje, y cuando el alumno asume que la palabra o estructura de la L1 existe en la L2, se verifica el fenómeno de la semejanza asumida en la producción (Ringbom, 2006). Sin embargo, principalmente en el caso de los falsos cognados, la semejanza formal puede inducir al error. De cualquier forma es necesario tener presente el hecho de que la semejanza funciona de manera diferente a lo largo del proceso de aprendizaje y varía según el modo de aprendizaje y características propias del estudiante.

La percepción que tienen los estudiantes de las semejanzas y diferencias entre la L2 y la L1 ha sido denominada psicotipología (*psicotypology*) (Kellerman, 1977, 1978, 1983). Estudios como los de Singleton (1987), Bild y Swain (1989) y Möhle (1989) confirmaron el efecto de la psicotipología. Con todo, se ha observado también que la percepción de la diferencia tipológica entre las dos lenguas implicadas cambia a medida que avanza el conocimiento del estudiante sobre la L2, dado que cuanto menos el estudiante sabe sobre la lengua meta, más está propenso a recurrir a sus conocimientos lingüísticos previos (Kellerman, 1983). De ahí que las investigaciones hayan encontrado resultados en los que se observa más transferencia de la L1 en niveles bajos de competencia.

Algunos estudiantes asumen que toda estructura o palabra que se parezca a la L1 debe ser incorrecta, pues piensan que la L2/LE tiene que ser necesariamente diferente de su L1. Esta percepción los induce, por un lado, a evitar el término que consideran semejante, y por otro, a elegir formas no esperadas en su intento de producir L2. Este fenómeno es motivado por la sensación de que los elementos lingüísticos de la L2 deben ser necesariamente diferentes de la L1 y se conoce por *homiofobia* (Kellerman, 2000), el miedo a la similitud.

Dado que la semejanza croslingüística está condicionada a la percepción que tiene el estudiante de esta característica, Ringbom (2006) destaca por lo menos cinco dimensiones diferentes de la semejanza que son relevantes al estudio de la transferencia: a) semejanza asumida *versus* semejanza real; b) semejanza, diferencia y relación cero en ítems y sistemas; c) semejanza de forma *versus* semejanza de función/significado; d) el efecto de la semejanza que afecta a la comprensión *versus* semejanza que afecta a la producción, y e) semejanza para transferencia de L1 *versus* transferencia de L2 en el aprendizaje de L3. El autor también resalta la diferencia de expectativas que tienen los estudiantes respecto a ítems y a sistemas, según las que los alumnos son conscientes de que los ítems son diferentes en L2 respecto a la L1, pero tienden a asumir que los procedimientos de los sistemas subyacentes son los mismos que en su L1, u otra lengua previamente estudiada, salvo en el caso de que hayan sido advertidos de lo contrario.

En cuanto a los tipos de relación que el estudiante establece entre los sistemas de las lenguas implicadas en situación de semejanza croslingüística, Ringbom (*op. cit.*) identifica: a) una relación de semejanza en la que el estudiante establece una conexión de uno para uno



con la otra unidad, normalmente en la L1; b) una relación de diferencia, en la que el estudiante percibe la similitud y la diferencia; y c) una relación cero, en la que el estudiante no es capaz de relacionar aspectos de la L2 que está estudiando a conocimientos previos, como por ejemplo en el caso de estudiantes ingleses de alemán cuando se enfrentan a los géneros gramaticales porque no tienen esa característica en su L1.

Además, el autor explica que hay tres niveles de transferencia lingüística: general, de ítem y de sistema. La transferencia en el nivel general se refiere a cuánta similitud croslingüística el estudiante es capaz de percibir entre ítems y sistemas en las dos lenguas, desde un alfabeto común hasta el número palabras cognadas, por ejemplo. La transferencia en el nivel de ítem se basa en la similitud subyacente de forma y una similitud asumida de función o significado entre la L1 y la L2. En lo que se refiere a la transferencia de sistema, el alumno asume que hay una semejanza subyacente de función, cuando no se verifica necesariamente una similitud formal. De todos modos, la distinción entre semejanza formal y funcional/semántica no es muy precisa.

Cuanto más próximas son las lenguas, más común es encontrar 'transferencias generales'. Ese tipo de transferencias tiene un efecto facilitador de la adquisición en actividades de comprensión, lo que resulta en transferencia positiva. Como la transferencia general se basa en los ítems y en algunas equivalencias funcionales entre los sistemas en cuestión, cuanto más ítems de la lengua meta son percibidos por los estudiantes como semejantes a los de lenguas que el estudiante ya conoce previamente, más transferencia general habrá.

En el caso de la 'transferencia de ítem', el efecto, así como en el caso de la transferencia general, es positivo, aunque algunas veces puede ser también negativo. Un ítem puede ser tanto palabras, ítems lexicales, como fonemas, morfemas o frases. Las semejanzas croslingüísticas que establezcan los estudiantes pueden o no llevar a errores. Los errores serán menos frecuentes en el ámbito de la comprensión una vez que el contexto tanto lingüístico como extra-lingüístico ayuda a eliminar los errores que son comunes en la producción.

Varios estudios indican que hay transferencia de una interlengua a otra, pero Gass y Selinker (2008) defienden el desarrollo de estudios longitudinales que demuestren no sólo que este

tipo de transferencia existe, sino también que no desaparece, es decir, que persiste a lo largo del proceso de aprendizaje.

## 2.5. Lenguas afines: el portugués y el español

Dado que en el presente trabajo se estudian los errores en las palabras léxicas que se observan en la producción escrita en español<sup>13</sup> por hablantes del portugués variedad brasileña, y con el objetivo de aportar más información sobre la especificidad del contexto de enseñanza de lenguas próximas, a continuación se exponen y se comentan algunas investigaciones realizadas en el ámbito del aprendizaje de lenguas en contextos de lenguas próximas, más concretamente sobre el aprendizaje de ELE por hablantes del portugués (L1).

Cuando se habla de lenguas próximas o lenguas afines, se hace referencia a la similitud tipológica, es decir, al hecho de que se trata de lenguas flexivas, que comparten muchos rasgos lingüísticos, aparte de que tienen una evolución histórica similar (Ainciburu, 2007). A partir de la aproximación a la bibliografía existente de trabajos realizados en el contexto de lenguas afines (ej.: Almeida Filho, 1995; Natel, 2002; Sánchez Iglesias, 2003; Calvi, 2004; Durão, 2004; Greco, 2006; Ainciburu, *op. cit.*), se verifica que los estudiantes recurren a su L1 de diferentes maneras. Al principio, en los niveles iniciales, los estudiantes generalmente recurren más a su L1 porque perciben que esta estrategia les facilita el proceso comunicativo. Sin embargo, a medida que van adquiriendo un mayor conocimiento en la lengua meta (L2), empiezan a darse cuenta de las diferencias entre las lenguas y ya no se apoyan en la transferencia como al principio de su proceso de aprendizaje.

La idea de que a menor distancia interlingüística, menor grado de interferencia (Lado, 1957) no se comprueba una vez que, en el caso de lenguas afines en contacto, la distancia entre ellas es tan pequeña que es muy difícil separar los dos sistemas, por lo que son frecuentes los solapamientos y saltos continuos a todos los niveles: fonológico, léxico y gramatical (Martín del Rey, 2004). No se puede negar la interferencia de la L1 en el caso de lenguas afines y la relevancia de la L1 en el aprendizaje de la L2 en el caso de lenguas próximas es indiscutible. Sin embargo, la interferencia puede generar resultados positivos o negativos.

---

<sup>13</sup> Cuando hablamos del aprendizaje del español nos referimos a la variedad culta estándar, tanto del español peninsular como el de las Américas.

La proximidad tiene un efecto positivo en el aprendizaje de lenguas muy afines en los niveles bajos, es decir, el aprendizaje es más rápido generalmente en las destrezas receptivas (Lenarduzzi, 1999). Con todo, es importante destacar que cuando el estudiante llega a un cierto nivel de competencia, se produce un estancamiento en el aprendizaje causado generalmente por las interferencias que impiden que la interlengua del estudiante se aproxime a la L2, lo que hace que la evolución sea muy lenta (Martín del Rey, 2004; Lenarduzzi, *op. cit.*). Para Martín del Rey (*op. cit.*: 87) “cada lengua presenta unas características particulares y que en el caso de lenguas afines no se pueden aplicar todas las teorías de aprendizaje, ya que por su idiosincrasia estas lenguas presentan y necesitan unas consideraciones determinadas”. Sin embargo, en lo que se refiere a las destrezas receptivas, “cuanto más afines sean las lenguas más fácil será la comunicación entre sus hablantes” y en el caso específico de interacción comunicativa entre un hablante de español (L1) y un hablante de portugués (L1), ésta puede ocurrir en su lengua materna porque “ambos entienden la lengua del otro” (Domínguez González, 2008: 27).

Estudios realizados para el proyecto EUROM4 (Blanche-Benveniste *et al.*, 1997) confirmaron que hay una gran similitud en la gramática, semántica e incluso pragmática entre cuatro de las lenguas romance: español, italiano, portugués y francés, lo que permite que un hablante de una de estas lenguas comparta por lo menos un conocimiento pasivo de las demás (Cresti, 2005).

En el caso específico de las lenguas que nos interesa estudiar, el español (L2) y el portugués (L1), la proximidad lingüística entre ellas se debe a que las dos lenguas además de tener el mismo origen románico, son semejantes tipológicamente y las bases culturales en las que se apoyan se comparten ampliamente (Israel, 2001). De acuerdo con la autora, todo eso implica que las dos lenguas comparten gran parte de la fuente del léxico, lo que facilita el proceso de comprensión del nuevo idioma en los primeros niveles de enseñanza. Sin embargo, la investigadora también resalta que muchas veces esa facilidad es un engaño, y afirma que es en este momento que se hace necesaria la intervención del profesor en la elucidación del significado y para que el proceso de enseñanza-aprendizaje del español por brasileños sea eficaz.

En esa misma línea, Jordan (1991) y Greco (2006) defienden la idea de que el aumento de la conciencia de los estudiantes sobre los aspectos lingüísticos que son diferentes entre las lenguas próximas en contacto (L1 y L2) es un instrumento extremadamente útil en el desarrollo de la competencia lingüística de los alumnos.

Las lenguas portuguesa y española, de acuerdo con Takeuchi (1984), son de las lenguas románicas las que presentan mayor transparencia lingüística entre sí. También Godoy (2001: 330) afirma que el portugués y el español son dos lenguas hermanas mutuamente inteligibles (aunque no simétricamente) y que la fosilización<sup>14</sup> de la interlengua estaría apoyada básicamente en la hipergeneralización (*overgeneralization*) interlingüística, es decir, en la proyección de las reglas marcadas de la L1 porque el estudiante adquiere la sensación de que todo ‘es lo mismo’.

Debido a la proximidad, es frecuente encontrar a brasileños que se consideran hablantes de español por el simple hecho de sentirse dominadores de unos pocos rasgos fonéticos o unidades léxicas (Eres, 2001). Como consecuencia de la confianza, unida a la facilidad de la intercomprensión, surge la manifestación lingüística conocida como ‘portuñol’. Este término se refiere a las producciones lingüísticas que son el resultado de una mezcla de elementos del portugués y del español (Campos, 2000). Según Lipski (2006) el portuñol/*portunhol* es un término que puede referirse a dos situaciones distintas. El autor explica que primeramente se ha usado para referirse al lenguaje de las comunidades en situaciones de frontera (comunidades localizadas entre Brasil, Uruguay, Paraguay, Argentina, por ejemplo). En segundo lugar, el término portuñol/*portunhol* se ha utilizado para describir la interferencia de la L1 que ocurre cuando un hablante de portugués intenta hablar español o viceversa, pero no es capaz de suprimir los rasgos de la L1.

Para Celada y Rodrigues (2005: 3), el ‘portuñol’ es una alternativa que las autoras designan como “efecto de una ilusión de competencia espontánea”. Eso se verifica principalmente en el aprendizaje espontáneo, cuando la afinidad hace que los hablantes recurran a la transferencia como estrategia comunicativa (Schmid, 1994). Se observa un rápido progreso de la interlengua, pero también se verifica una tendencia a la fosilización, y fenómenos de

---

<sup>14</sup> La fosilización es un concepto que se refiere a la estagnación del desarrollo de la interlengua (Van Patten y Benati, 2010).

contaminación entre los sistemas lingüísticos en contacto (Meo Zilio, 1989; Cancellier, 1996).

Sobre el tema de la similitud entre las lenguas portuguesa y española, se pueden destacar dos investigaciones que tratan de la comprensión oral y escrita del español por hablantes nativos del portugués. El estudio de Jensen (1989), por un lado, constata que los hablantes de portugués tienen un buen nivel de comprensión del español hablado (50% - 60%). Por otro lado, el estudio de Henriques (2000) expone que la comprensión lectora de textos escritos en lengua española por parte de brasileños puede llegar a tener un alto grado de comprensión léxico-semántica, de 58% a 94%. En este estudio, se seleccionaron 300 hablantes nativos de portugués y 300 hablantes nativos de español que nunca habían estudiado la otra lengua. Los sujetos hablantes de lengua portuguesa eran estudiantes brasileños del primer año de los cursos universitarios de Biblioteconomía, Nutrición, Farmacia, Medicina, Economía y Geografía de los estados de Rio de Janeiro y São Paulo, con edad media entre 18 y 21 años, sin ningún tipo de conocimiento formal o informal de lengua española. Según Henriques (2000: 4), los falsos cognados, las expresiones idiomáticas y los articuladores son los elementos principales, que si no se conocen en cada una de las lenguas, puede comprometer el sentido y la comprensión de todo el texto y afirma que:

*Com base nos resultados acima, podemos afirmar que, entre português e espanhol, a transferência de conhecimento, idealmente, pode chegar a mais de 90%. Todo o conhecimento que o aprendiz/falante transfere, com acerto, da sua língua nativa para a segunda língua (ou língua estrangeira) facilita o processo de aquisição/aprendizagem e a compreensão (oral e/ou escrita).*

Silva (2002), en su trabajo sobre las dificultades o bloqueos del estudiante en el nivel léxico-semántico por la interferencia de la lengua materna (L1, portugués) en el aprendizaje de una lengua extranjera (LE, español), observa que la mayor o menor facilidad en el aprendizaje del español está asociada a una mayor o menor transferencia del léxico, debido a la gran cantidad de palabras cognadas entre las dos lenguas, además de los incontables falsos cognados. La autora también afirma que la proximidad entre ambas lenguas no disminuye el grado de dificultad en el proceso de adquisición-aprendizaje del español porque se deben tener en cuenta los diferentes elementos que se combinan y se oponen en la aparente similitud estructural de estos dos sistemas lingüísticos.

Natel (2002), en su artículo 'La proximidad entre el portugués y el español ¿facilita o dificulta el aprendizaje?', enfatiza la importancia que tiene la lengua materna (L1) en el aprendizaje de una lengua extranjera (L2), pues en el caso de lenguas muy próximas como es el del español y el portugués, la facilidad de comprensión desencadena una serie de comportamientos como, por ejemplo, el hecho de que los alumnos empiezan a arriesgarse en la producción de la lengua meta desde los primeros días de clase. Además, va más allá al afirmar que el estudio de la proximidad influye en la pedagogía utilizada en el aula, y señala que no se puede seguir los mismos procedimientos que son utilizados en una clase de inglés, por ejemplo, poniendo de relieve la necesidad de tener en cuenta las cuestiones metodológicas específicas relacionadas con la enseñanza del español a brasileños.

Merinero (2002) ha trabajado la reflexión sobre los conceptos de facilidad, un concepto de aprendizaje subjetivo, y similitud, un concepto lingüístico, no subjetivo. Además, también ha reflexionado sobre las relaciones derivadas de ese supuesto binomio para llamar la atención sobre el trabajo con las ideas generalizadas sobre las lenguas, la reducción de una lengua a un conjunto de reglas fijas y estáticas donde el uso y las variantes no son relevantes (en el caso del español, desde el hecho de ser una lengua idéntica pero con diferente pronunciación hasta la idea de lengua tramposa). En el caso de aprendices lusohablantes este trabajo cobra un matiz especial por la creencia extendida de la existencia del fenómeno conocido como 'portuñol'.

La proximidad lingüística entre las referidas lenguas ha sido tratada en el artículo de Fialho (2005) con especial atención al tema del vocabulario heterosemántico, es decir, los falsos amigos o falsos cognados, palabras con semejanza en la forma, pero con diferente significado. Debido a la cercanía tipológica existente entre los dos idiomas, se considera que el profesor que enseña el español a hablantes de portugués deberá esforzarse para que sus alumnos superen esta cercanía entre los dos idiomas que, además de fomentar las transferencias, puede favorecer la fosilización. Briones (2000) comparte esa preocupación y explica que los errores más comunes en la traducción entre el portugués y el español, y también entre el español y el portugués, se deben al desconocimiento de los heterosemánticos. Andrade Neta (2000) con todo, amplía el campo que puede suponer dificultad para los brasileños estudiantes de ELE, y afirma que los errores en el nivel léxico pueden ocurrir no sólo por causa de los vocablos heterosemánticos, sino también por los

heterotónicos (forma similar, sílaba tónica diferente) y heterogénicos (forma similar, género diferente), porque son palabras que comparten rasgos en el campo ortográfico, semántico o fonológico y que por ello los estudiantes pueden incurrir en una falsa sensación de conocimiento de la forma de la lengua meta.

De acuerdo con Schmitz (1970), los contrastes léxicos en español y en portugués pueden organizarse en cuatro categorías básicas: a) ítems léxicos idénticos en ambas lenguas, pero con diferencia de significado en las dos; b) ítems léxicos diferentes en ambas lenguas, pero con el mismo significado en las dos lenguas; c) el ámbito del ítem léxico español es más amplio que la palabra correspondiente en portugués, y d) el ámbito del ítem léxico en portugués es más amplio que el correspondiente en español.

Según Camorlinga (1997), el aspecto que más revela la similitud entre las dos lenguas se constata en el léxico, ya que 85% de las palabras tienen un origen común. Más concretamente, Almeida Filho (2001) afirma que el léxico entre las dos lenguas está constituido por 60% de cognados idénticos y 30% de cognados falsos. Con todo, Fialho (2005) pone énfasis en los aspectos positivos que tiene la similitud entre las lenguas portuguesa y española, ya que permite una buena comunicación entre alumnos y profesor desde el primer día de clase.

Al tratar el tema de la proximidad lingüística desde la perspectiva de los brasileños, González (2000) explica que, de manera general, se verifican al menos tres escenarios posibles: a) la proximidad se caracteriza como elemento facilitador; b) la proximidad lingüística y la facilidad psicológica no se confirman, y c) la construcción de una lengua exótica en un intento de alejarse del portugués. Este punto de vista nos recuerda la idea de Kellerman (1978) que afirma que hay que tener en cuenta el hecho de que la proximidad lingüística no siempre es la misma que siente el hablante al estar en contacto con la LE, es decir, la cercanía psicológica no siempre corresponde a la lingüística.

En cuanto a las creencias de los alumnos universitarios brasileños en una situación de aprendizaje de lengua española, Soto (2002) explica que no sería absurdo afirmar que una gran parte de los alumnos que empiezan a estudiar en la universidad demuestran tener una representación esquemática de la lengua española que se basa en la creencia de una relación de proximidad y transparencia entre las lenguas portuguesa y española. Además, la

autora afirma aún que esta idea de cercanía y transparencia viene muchas veces reforzada por el hecho de que muchos de esos alumnos aprobaron el *vestibular*<sup>15</sup> en lengua española estudiando solamente algunos meses antes del examen.

Por todo ello, Barros y Costa (2011: 108) defienden la utilización de materiales didácticos que tengan en cuenta la especificidad de la enseñanza de ELE a alumnos lusohablantes. Según los autores, estos materiales deben considerar la proximidad lingüística de manera constructiva, es decir, aprovechar el conocimiento previo que el alumno tiene de su L1 por un lado, y por otro insistir más en aquellas estructuras y elementos lingüísticos que son diferentes en ambas lenguas (que no se restringen sólo a una lista de palabras heterosemánticas y heterogénicas). Así, será posible establecer los contrastes necesarios que ayudarán a los alumnos a superar la falsa percepción de la facilidad de aprendizaje de la lengua española y la creencia de que no es necesario estudiarla porque las lenguas son muy parecidas.

Aunque muchas veces los errores que se observan en la interlengua de estudiantes lusohablantes de ELE se tratan de errores de corrección y que rara vez afectan la comunicación como señala Camorlinga (1997), el autor afirma que si se pretende usar la lengua española de forma precisa y profesional, su aprendizaje se convierte en un proceso largo y difícil. Asimismo, Espiga (1999) destaca que la situación peculiar del aprendizaje de español por hablantes de portugués, pese a que se caracterice por ser un proceso corto en distancia, parece exigir más tiempo y dedicación.

Como se ha visto, la proximidad lingüística entre el español (L2) y el portugués (L1) influye en el proceso de aprendizaje de ELE de los estudiantes brasileños de varias maneras. Por un lado, vimos que en un contexto de lenguas afines, la cercanía acelera el aprendizaje principalmente en el caso de las destrezas receptivas y en los niveles iniciales de competencia. Por otro, los estudiantes pueden desarrollar estrategias como la evitación, causada por el miedo a la similitud (*homiofobia*), cuando perciben la semejanza como algo a ser evitado. También pueden generar muestras de interlengua con interferencias causadas por la hipergeneralización de reglas, cuando perciben las lenguas como muy similares y se apoyan totalmente en sus conocimientos previos. Todos esos factores lingüísticos y

---

<sup>15</sup> Examen de acceso a la universidad.



psicológicos afectan al proceso de aprendizaje del español por estudiantes brasileños porque generan comportamientos y actitudes que pueden interferir tanto de forma positiva como negativa en la adquisición de la L2.



### 3. EL ERROR

Dado que el objetivo central de esta tesis es estudiar los errores léxicos producidos en un contexto de lenguas próximas (L1 portugués y L2 español), el presente capítulo procura reunir algunas de las principales apreciaciones existentes con respecto al error en el ámbito de la Adquisición de Segundas Lenguas (ASL). Consideramos también, así como Fernández (1997: 19), que los errores son “índices de los estadios por los que el aprendiz atraviesa en el camino de apropiación de la lengua meta”.

Antes de pasar al apartado siguiente, nos parece pertinente comentar que aunque nos centramos en el estudio de los errores, comprendemos que la interlengua de los estudiantes también se compone de formas correctas. Sin embargo, estamos de acuerdo con Alexopoulou (2006) cuando explica que al focalizar en los errores se puede conseguir información sobre las características de la interlengua de los alumnos y conocer un poco más sobre algunas de las dificultades que experimentan a lo largo de su proceso de aprendizaje de una segunda lengua (L2).

#### 3.1. La identificación, descripción y clasificación del error

No cabe duda de que el error es un elemento que siempre ha hecho parte del proceso de aprendizaje de segundas lenguas. De manera general, suele haber dos actitudes contrarias ante el concepto de error: una negativa y una positiva. La actitud negativa, por un lado, ve el error como un elemento que afecta al proceso de aprendizaje y aleja al estudiante del nivel deseado de competencia en L2. Por otro, la actitud positiva ante el error lo considera parte del proceso de aprendizaje e indicación de las distintas etapas de aprendizaje que atraviesa el estudiante (Santos Gargallo, 1993). Sin embargo, la percepción del error ha cambiado a lo largo del tiempo.

Como se ha visto en el capítulo 1 (Los paradigmas de estudio de la Adquisición de Segundas Lenguas), algunos investigadores habían percibido que la comparación entre los sistemas lingüísticos de la L1 y L2 en un intento de prever y evitar los errores (Análisis Contrastivo) no era suficiente para explicar los errores de los alumnos, porque no tenían en cuenta el proceso creativo que realizan los estudiantes y que se refleja en su interlengua. Así, los

investigadores pasaron a estudiar los errores que producían los estudiantes y los comparaban con la forma correspondiente de la lengua meta (L2) (Análisis de Errores).

El principal responsable de esa nueva percepción del error es el investigador Pit Corder que identificó en los errores de los estudiantes de L2 una importante fuente de información. Según Corder (1967, 1981), los errores a) indican al profesor cuánto ha progresado el alumno, b) proporcionan al investigador información sobre cómo se aprende una L2, y también c) ayudan a los alumnos a reformular sus hipótesis sobre la lengua que están estudiando. Desde el punto de vista de la didáctica, la información que se obtiene a través de los errores de los alumnos le permite al profesor replantear su práctica docente en cada momento. Por eso, es importante que el profesor sea capaz tanto de detectar y describir los errores lingüísticamente, como también de comprender las razones psicolingüísticas que llevaron al alumno a producir el error.

Corder (1981) distingue entre los términos 'error' y 'falta o equivocación' (*error and mistake*). La diferencia entre los errores (errores de competencia) y las faltas (errores de actuación) está en la capacidad del estudiante de autocorregirlos. Si el estudiante es capaz de identificar y proporcionar la forma correcta, la desviación se considera una falta, si no, se considera un error. Dulay *et al.* (1982), a su vez, utilizan el término *error* para referirse a cualquier desvío con relación a la lengua meta.

Con todo, James (1998) prefiere la distinción que hace Corder y afirma que la competencia es el conocimiento de la regla por parte del alumno, y que un *error de competencia* es cuando uno no se da cuenta del error y no puede autocorregirse. Por otro lado, la *equivocación* está relacionada con el desempeño y el alumno puede autocorregirse porque tiene el conocimiento necesario para identificar el desvío y repararlo. Así, James (*op. cit.*) utiliza el término 'desviación' para cualquier forma que no se corresponde a la L2 y, tras hacer un repaso de varias nomenclaturas, sugiere clasificar las desviaciones en cuatro tipos:

- Lapsos (*slips*): pueden ser detectados y corregidos por el autor sin ayuda;
- Equivocaciones (*mistakes*): sólo pueden ser corregidas por sus autores si alguien las señala. Si sólo el hecho de apuntar el problema es suficiente para la autocorrección, entonces se considera una equivocación de primera orden. Pero,

si se necesita más información, entonces, se tiene una equivocación de segundo orden.

- Errores: son aquellas desviaciones de la lengua meta que no se pueden autocorregir hasta que no se administre *input* relevante y que éste se haya convertido en *intake*<sup>16</sup> por el estudiante.
- Solecismos: son infracciones de las reglas de corrección que suelen generar conflictos con las intuiciones de los hablantes nativos.

Sin embargo, definir qué es un error no es tarea fácil (James, 1998; Ellis, 2008). Eso es así porque, por un lado está la gramaticalidad y, por otro, la aceptabilidad, lo que hace que un ítem pueda ser correcto gramaticalmente, aunque su uso puede no ser adecuado.

Corder (1967, 1981) estableció las bases de la investigación del modelo de Análisis de Errores y definió tres etapas esenciales de su realización: a) identificación del error, b) descripción del error y c) explicación del error, que explicamos a continuación.

### 3.1.1. Identificación del error

El primer paso en la realización de un análisis de los errores es detectar los errores. Cuando se detecta un error, uno se da cuenta de su existencia y nada más. Sin embargo, hay factores que influyen en la detección del error, como el medio (oral o escrito), por ejemplo. Esto es así porque se cree que es más difícil detectar un error en la lengua hablada en un contexto informal que en la lengua escrita en un contexto formal. Además, detectar errores en la pantalla también parece ser más difícil que en la versión impresa y apuntar el error de uno mismo es más difícil que apuntar los errores de los demás.

Asimismo, tras identificar el error, no siempre es fácil delimitarlo. Muchas veces la sentencia no contiene un error, sino que es una sentencia defectuosa. Pueden ser errores cometidos o errores de omisión. Como se ha visto anteriormente (véase nota al pie 4, p. 18), los errores de omisión se deben a que algunos estudiantes adoptan la estrategia de la evitación principalmente en contextos de lenguas afines, en los que las lenguas implicadas comparten

---

<sup>16</sup> El término *intake* fue utilizado por Corder en 1967 para marcar la diferencia entre las muestras de lengua a las que el estudiante está expuesto (*input*) y la información lingüística que el estudiante realmente asimila (*intake*) (VanPatten y Benati, 2010). De acuerdo con el Diccionario de Términos Clave de ELE (Centro Virtual Cervantes, 2012), el término *intake* es también conocido como “aducto” o “caudal lingüístico” en español.

una serie de características. Muchos investigadores sugieren que además de tener atención en las formas erróneas que producen los estudiantes, es importante observar también lo que dejaron de producir (Schachter, 1974). Si se ignora lo que dejan de producir los estudiantes, es probable que la interpretación de los resultados no sea la más adecuada, porque aunque un determinado grupo de estudiantes produzca menos errores que otro respecto a un determinado ítem, es importante averiguar si esos estudiantes emplearon el ítem estudiado en sus producciones de forma correcta o si lo evitaron completamente.

Con todo, cuando la identificación del error se hace con fines de evaluación, es necesario tener en cuenta que los errores tienen características que influyen en la irritabilidad del evaluador. James (1998) distingue tres factores: a) si el error es predecible (*predictability*); b) la relación social de los interlocutores, y c) el grado de infracción de las normas sociales.

### 3.1.2. Descripción del error

En esta fase, se lleva a cabo una actividad de análisis lingüística que procura describir las características de los errores. Del mismo modo que es difícil identificar el error, también es difícil describirlo. Después de la identificación del error, la reconstrucción de la sentencia en su contexto no siempre es posible. Aunque el alumno esté disponible para consulta, la descripción del error va a depender de la reconstrucción elegida (Ellis, 2008).

De acuerdo con James (1998), los errores pueden pertenecer a distintos niveles según la modalidad, que puede ser productiva (codificación) o receptiva (decodificación); según el medio, oral o escrito; y según el nivel, o sea, si está en la sustancia, en el texto o en el discurso. Los errores en el ámbito de la sustancia pueden ser grafológicos (errores de puntuación, errores tipográficos, errores disléxicos y errores confusos o fonológicos) o fonológicos (errores de pronunciación o errores de codificación escrita). Los errores a nivel de texto pueden ser de léxico o de gramática y en el nivel del discurso son los errores de coherencia. En la figura 3 abajo se puede apreciar más claramente un esquema de esa propuesta de clasificación.

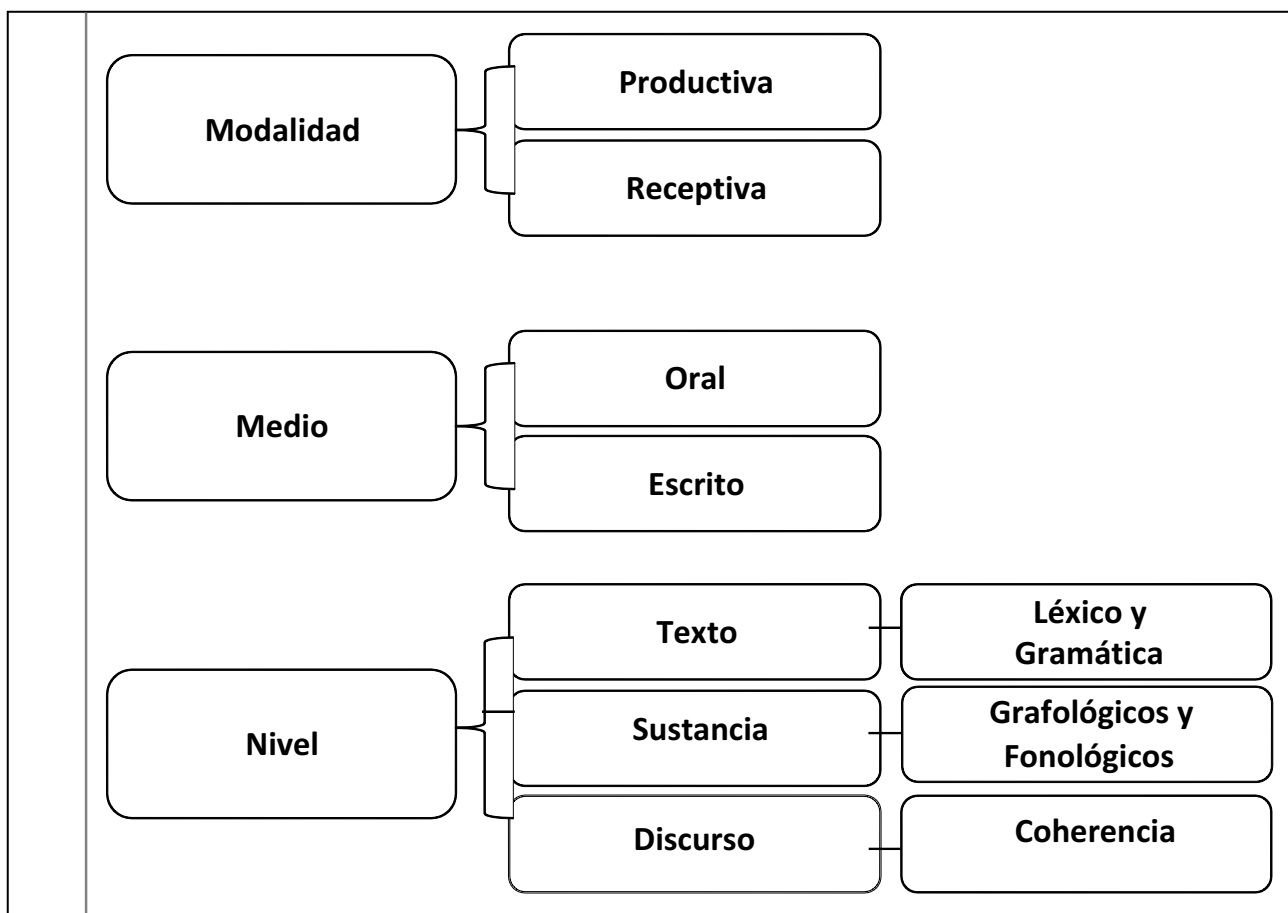


Figura 3. Esquema de clasificación de los errores de James (1998).

En el ámbito de las lenguas próximas, Sánchez Iglesias (2003: 85) sigue la taxonomía lingüística (fonología, morfología, sintaxis, léxico, semántica, pragmática y ortografía) y describe los errores según el nivel lingüístico en el que ocurre el error, distinguiendo seis niveles: fonético-ortográfico, léxico, morfológico, morfosintáctico, sintáctico y semántico.

### 3.1.3. Clasificación del error

Después de detectar y describir el error, el paso siguiente suele ser la categorización de los errores. Para hacer una clasificación de los errores, primero es necesario diseñar un instrumento metodológico que comprenda las categorías de los errores que se pretende estudiar.

Una de las razones por la que se clasifican los errores es para que se pueda contar el número de veces que determinada forma aparece en el texto. Sin embargo, Meara (1984)

enfatisa que estas taxonomías no representan más que un primer paso y no son suficientes para explicar qué adquieren los estudiantes, ni cómo o cuándo lo hacen.

Agustín Llach (2011b: 24) hace un balance de algunas de las críticas y ventajas que tiene la clasificación de los errores léxicos y comenta que

se han señalado algunas limitaciones de las taxonomías de errores léxicos como su subjetividad clasificatoria, su carácter restringido pues se limitan a tratar el error, cuando buena parte de las conductas léxicas no desembocan en un error, su escaso poder predictivo *a priori*. Con todo ello y pese a estas limitaciones, los diversos estudios han puesto de manifiesto la gran utilidad del análisis de errores como metodología de investigación en los procesos de adquisición y enseñanza del léxico en la lengua extranjera.

Dada la variedad, extensión y especificidad de las diferentes taxonomías que se han desarrollado con el objetivo de profundizar en el análisis de los errores léxicos, recomendamos, para una revisión de las clasificaciones más representativas utilizadas en ASL, consultar la investigación sobre los errores léxicos y la corrección en textos escritos en inglés como lengua extranjera de Agustín Llach (2011a: 76).

Con todo, una clasificación de los errores léxicos ampliamente empleada, es la que se basa en la dicotomía formal/semántico, según el tipo del error (James, 1998; Ringbom, 2001). La principal justificativa de esa distinción se debe a la organización del lexicón mental, cuyas investigaciones sugieren que se organiza siguiendo tanto principios formales como semánticos.

De acuerdo con James (*op. cit.*), los errores formales son aquellos que afectan al significante, se consideran menos graves que los errores semánticos dado que no suelen interrumpir la comunicación y se clasifican en 'selección equivocada de la forma', 'forma equivocada' y 'distorsiones'. Por otro lado, los errores semánticos son aquellos que afectan al significado de la palabra, por lo que son más propensos a interferir en la transmisión del mensaje.

En cuanto a los errores de 'selección equivocada de la forma', se tratan de pares de palabras que se parecen y suenan de forma similar (*synforms*) (Laufer, 1992). Las *synforms* presentan mismo número de sílabas, misma sílaba tónica, hacen parte de la misma clase de palabras, empiezan de la misma forma, tienen algunos fonemas en común y tiene fonemas que comparten características similares. Se dan afinidades en este tipo de error entre nativos y estudiantes de L2, lo que no ocurre en otras categorías. Estos errores son comunes cuando



la L1 y la L2 son lenguas afines, es decir, lenguas en las que se observa un alto grado de semejanzas. Como se ha visto anteriormente, la semejanza implica a veces en una dificultad creada justamente por la similitud interlingüística, cuando lo esperado en estos casos sería que la semejanza facilitara el proceso de aprendizaje de la L2. Sin embargo, James (1998) afirma que en el uso receptivo de la lengua, el contexto neutraliza la mayor parte de las asociaciones semánticas erróneas, y por lo tanto, en este caso la semejanza croslingüística actúa como una característica positiva que facilita la comprensión de la lengua meta.

Los 'errores de forma equivocada' son las palabras que no existen en la lengua meta (L2). Pueden tener origen en la L1, interferencia interlingüística, o en la L2, interferencia intralingüística: a) si la palabra de la L1 es usada en una producción en L2, hablamos de préstamo (*borrowing*); b) si la nueva palabra derivada de la L1 sufre algún tipo de adaptación a la L2 tenemos una acuñación (*coinage*); c) si la palabra en la L2 es el resultado de una traducción literal de una palabra de la L1, tenemos un calco (*calque*).

Las 'distorsiones' son errores intralingüísticos que originan formas no existentes en la lengua meta. Son el resultado de una aplicación errónea de una de las cuatro operaciones siguientes: omisión, inclusión innecesaria, selección equivocada, ordenamiento equivocado y mezcla.

En lo que se refiere a los errores semánticos, tal y como se ha dicho anteriormente, son aquellos errores que afectan al significado. James (*op. cit.*) los clasifica en 'errores de confusión de relaciones de sentido' y 'errores de colocación'.

Los 'errores de confusión de relaciones de sentido' se deben a que hay una considerable evidencia neurolingüística que sugiere que los humanos guardamos palabras en el lexicón mental en términos de relaciones de sentido. Por ello, el autor cree razonable categorizar los errores léxicos con referencia a ese sistema: a) usando un término general cuando uno más específico es necesario (hiperónimo por hipónimo); b) usando un término muy específico (hipónimo por hiperónimo); c) usando el menos apto de dos co-hipónimos, y d) usando el término erróneo de un conjunto de sinónimos.

Los 'errores de colocación' son errores que ocurren generalmente por desconocimiento del término en la L2 o falta de conocimiento declarativo. Cuando el estudiante desconoce el

término en la L2 y lo toma prestado de la L1, la consecuencia es un error de transferencia de la L1. No obstante, cuando el estudiante conoce el término en la L2, pero falla al acceder a él y accede a un término de la L1 sustituto, tenemos un caso de una equivocación por interferencia de la L1.

Asimismo, Ringbom (2001) presenta una clasificación de los errores formada por cinco categorías: a) alternancia de código (*language switches*); b) acuñación (*coinages*); c) falsos cognados (*false friends*); d) calcos (*calques*), y e) extensiones semánticas (*semantic extensions*). Los calcos y las extensiones semánticas son las categorías que están relacionadas con la transferencia semántica.

Con relación específicamente a los trabajos que se han dedicado al análisis de errores en el ámbito de ELE, se puede decir que los trabajos de Vázquez (1991), Santos Gargallo (1993) y Fernández (1997) son los “más representativos” (Penadés Martínez, 2003: 2) y en los que se han basado muchas de las investigaciones que les sucedieron. A continuación presentamos de manera sucinta cada uno de estos modelos de clasificación.

Los criterios de clasificación de los errores que utilizó Vázquez (1991: 31-85) son los siguientes:

- Criterio lingüístico: errores de adición, omisión, selección falsa, colocación falsa, yuxtaposición;
- Criterio etiológico: errores intralinguales, interlinguales, de simplificación;
- Criterio comunicativo: errores de ambigüedad, irritantes, estigmatizantes;
- Criterio pedagógico: errores inducidos vs. creativos, errores transitorios vs. permanentes (fossilizados vs. fossilizables), errores individuales vs. colectivos, errores en la producción escrita vs. errores en la producción oral.

Santos Gargallo (1993: 91-97) propone la siguiente clasificación de los errores en su trabajo:

- Criterio descriptivo: errores de omisión, de adición, de formación errónea o de ausencia de orden oracional;
- Criterio pedagógico: errores de actuación y errores de competencia;
- Criterio etiológico-lingüístico: errores interlingüísticos e intralingüísticos;

- Criterio gramatical: fonológicos, ortográficos, morfológicos, sintácticos, léxicos, semánticos y pragmáticos;
- Criterio comunicativo: locales (cuando afectan a un elemento de la oración) o globales (cuando afectan a toda la oración).

Finalmente, la última de las tres propuestas de análisis de errores que hemos destacado es la de Fernández (1997: 44-48):

- Descripción lingüística: subsistemas: fonología, léxico, morfosintaxis, discurso;
- Descripción de estrategias: omisión, adición, falsa elección, falsa colocación;
- Criterio pedagógico: errores colectivos, errores individuales, errores transitorios, errores fosilizables, errores fosilizados, errores inducidos por la metodología;
- Explicación de los errores: errores ambientales, errores internos;
- Criterio comparativo: errores evolutivos, errores interlingüales, errores ambiguos;
- Efectos comunicativos: errores que distorsionan la comunicación.

De acuerdo con Penadés Martínez (2003: 20)

estas clasificaciones, en principio, no han estado guiadas por una pretensión de generalidad, al contrario, han surgido para examinar unos *córpore* concretos, y las peculiaridades de cada uno de éstos pueden haber determinado la organización de las clasificaciones con que se intentan explicar los errores en ellos contenidos, y, por otra parte, en modo alguno es una tarea fácil establecer una clasificación de errores lingüísticos, vista la situación de falta de acuerdo absoluta, en la lingüística teórica, acerca de una teoría gramatical que dé cuenta de las unidades gramaticales y de sus relaciones.

Como se ve, pese a las diferentes propuestas de clasificación que se han ofrecido, todavía no hay acuerdo entre las distintas taxonomías establecidas por los investigadores que se han dedicado a estudiar los errores en la producción de estudiantes de L2. Además, estas clasificaciones se han creado para atender las necesidades específicas de cada uno de estos estudios. Con todo, se observan correspondencias entre varios de los criterios utilizados por estas autoras para clasificar los errores. Para ilustrarlo, destacamos aquí la correspondencia que Penadés Martínez (*op. cit.*) establece específicamente sobre el análisis del origen de los errores: criterio etiológico (Vázquez, 1991), criterio etiológico-lingüístico y gramatical (Santos Gargallo, 1993) y criterio de la descripción lingüística (Fernández, 1997). Así, vemos que la identificación y el análisis del origen de los errores es una preocupación recurrente en las investigaciones en ASL, aunque se refieran al fenómeno de manera distinta.

En un estudio reciente sobre el aprendizaje de ELE por hablantes griegos, Palapanidi (2010) realiza un análisis de errores con base en el tipo del error léxico (formal o semántico) y el origen de esos errores (interlingüístico o intralingüístico). Como se ha visto más arriba, la idea de clasificar y analizar los errores según el tipo, proporciona información referente a la organización del lexicón mental. Por otro lado, el análisis de los errores tanto formales como semánticos (James, 1998), siguiendo el criterio etiológico (Vázquez, 1991; Santos Gargallo, 1993; Fernández, 1997), permite conocer si estos tipos de errores se deben a la transferencia (error interlingüístico) o al propio proceso de desarrollo de la interlengua (error intralingüístico).

#### 3.1.4. Explicación del error

La explicación de los errores intenta establecer la fuente del error. Para Ellis (2008), ése es el principal estadio en la investigación en ASL porque está relacionado con el intento de establecer el proceso responsable por la adquisición en segundas lenguas.

Los errores pueden deberse a diversas causas, es decir, pueden ser de origen psicolingüístico, sociolingüístico, cognitivo o pueden residir en la estructura del discurso. Sin embargo, la investigación en ASL se ha centrado sobre todo en el ámbito psicolingüístico. Generalmente, las clasificaciones de los errores según el origen psicolingüístico distinguen entre errores que tienen su origen en la L1 del hablante o en la lengua que están estudiando (L2). En esta línea, Richards (1971) distingue tres causas diferentes de los errores psicolingüísticos: a) errores de interferencia (*interlingual errors*), son los errores que tienen como causa la transferencia de la lengua materna o primera lengua (L1) que resulta en transferencia negativa; b) errores intralingüales (*intralingual errors*), son errores causados por las características propias de la L2, comunes a hablantes de cualquier L1; y c) errores de desarrollo (*developmental errors*), son aquellos errores propios del aprendizaje de la lengua y se identifican tanto en el aprendizaje de la L1 como de la L2.

Dulay y Burt (1974), a su vez, proponen una clasificación de los errores en: a) errores de interferencia; b) errores de desarrollo, y b) errores ambiguos o únicos, es decir, errores que no se pueden identificar su causa de manera clara. Las autoras consideran los errores únicos aquéllos en los que no se identifican un origen en la L1 ni en el proceso de desarrollo de la

interlengua. Esta propuesta ha recibido diversas críticas, dado que no engloba todas las posibles causas de los errores de los estudiantes de una L2. Un ejemplo de ello es la identificación del origen de algunos de los errores clasificados como ambiguos. Se ha averiguado que algunos de estos errores tenían su origen en la instrucción, por lo que Fisiak (1981) incorporó el término errores pedagógicos (*pedagogical errors*) para referirse a ellos. Estos errores pueden haberse originado por la influencia de la enseñanza o del material utilizado en las clases, es decir, son causados por factores extralingüísticos.

Schachter y Celce-Murcia (1977) sugieren una clasificación más general en la que se identifiquen los errores de transferencia (*transfer errors*) y los errores intralingüísticos (*intralingual errors*). Con todo, las autoras advierten que identificar el origen del error no es tarea fácil y que el investigador además de ser muy prudente cuando cree haber identificado la causa del error, debe siempre tener en cuenta el contexto en el que se ha detectado.

De manera general, las investigaciones en ASL han verificado que, en los estadios iniciales de aprendizaje de una L2, el 50% de los errores son de origen interlingüístico, es decir, se deben a la interferencia de la L1. Esos errores son más comunes en los niveles fonológico y léxico que en el gramatical y son más comunes en los estudiantes adultos que en los niños. Además, varios estudios revelan que la influencia de la L1 en el léxico disminuye al aumentar el dominio lingüístico de los estudiantes (ej.: Celaya y Torras, 2001; Navés *et al.*, 2005; Celaya, 2007). Asimismo, esos estudios también señalaron que en los estadios iniciales de adquisición se encuentran más préstamos<sup>17</sup> en la producción de los estudiantes, mientras que en los niveles avanzados el tipo de transferencia léxica más común son los calcos<sup>18</sup>.

Sin embargo, al contrario de lo que predicaba el Análisis Contrastivo, algunos estudios encontraron que la mayoría de los errores cometidos por los estudiantes de L2 tenían su origen en el propio proceso de aprendizaje de la lengua meta. Un ejemplo de ello es la

---

<sup>17</sup> El término “préstamo lingüístico” se refiere a una palabra, modificada o no, de una lengua que fue utilizada en la producción (oral o escrita) en otra lengua (Appel y Muysken, 1996).

<sup>18</sup> Según el *DRAE* (2001), un calco es la “adopción del contenido semántico de una palabra o expresión extranjera, traduciendo su significado mediante unidades lingüísticas propias de la lengua de recepción”. Un ejemplo de ello es el uso de la palabra “ratón” como herramienta en el ámbito computacional, traducción directa de la palabra del inglés “*mouse*”.

investigación que ha sido conducida por Dulay y Burt (1974) que, tras analizar 513 errores producidos por niños españoles estudiantes de inglés, identificaron que menos del 5% del total de errores se daban por interferencia de la L1. En este caso, se concluye que son las características propias del sistema lingüístico de la L2 lo que dificulta el proceso de aprendizaje (Dulay *et al.*, 1982) y que la causa principal de las producciones erróneas de los estudiantes de L2 no se debe ni a la influencia de la L1 ni de los conocimientos previos.

Asimismo, Ellis (2008) recoge una amplia lista de algunos de los principales resultados de los estudios empíricos en ASL que confirman la idea de que la mayoría de los errores en la producción en L2 son de origen intralingual. Más recientemente, Palapanidi (2010) en su análisis de los errores léxicos identificados en los textos escritos por estudiantes griegos de español, también encontró que la mayor parte de ellos tenían un origen intralingüístico, es decir, en el proceso de desarrollo de la interlengua.

Con todo, esos estudios se han realizado en contextos de aprendizaje de lenguas tipológicamente diferentes. De acuerdo con Ringbom (1986), en su estudio sobre el aprendizaje del inglés como lengua extranjera en Finlandia por sujetos hablantes de finlandés y de sueco, la distancia entre la L1 y la L2 está fuertemente relacionada con el índice de frecuencia de la interferencia. Según el autor, la interferencia se da principalmente en los primeros niveles de aprendizaje porque el estudiante cree que las funciones básicas de la lengua son idénticas a las de su L1, con diferencias únicamente en el léxico. Sin embargo, como se ha visto en el apartado 2.5. (Lenguas afines: el portugués y el español), en el caso del aprendizaje de español por lusohablantes, incluso una gran parte del léxico es común a ambas lenguas, lo que aumenta la probabilidad de que la interlengua del estudiante presente más interferencias originadas por su L1.

En un estudio sobre los errores léxicos en un contexto de lenguas afines, Sánchez Iglesias (2003: 90) optó por explicar los errores según su origen (interlingüístico o intralingüístico). El autor se apoya en la propuesta de Bustos Gisbert (1998) por considerar que éste presta especial atención a la transferencia lingüística, y utiliza la siguiente taxonomía para explicar los errores (Sánchez Iglesias, *op. cit.*: 99): a) interlingüísticos: acumulación, confusión, hipótesis falsa, regularización, simplificación, sobregeneralización; y los errores intralingüísticos: interferencia directa, por bifurcación, por colisión, por convergencia.

A partir de todo lo dicho, se constata que la identificación del origen del error no es una tarea fácil. Lo que es interesante destacar es que en base a las clasificaciones descritas anteriormente, los errores intralingüísticos y de desarrollo son igual de importantes que los errores provocados por la interferencia de la L1 porque proporcionan información sobre las estrategias e hipótesis de aprendizaje que usan los estudiantes durante del proceso de aprendizaje de la L2.

### 3.2. El error léxico

Aunque los análisis de la interlengua de los estudiantes de segundas lenguas revelen que los errores léxicos sean más numerosos que los errores gramaticales (Agustín Llach, 2011a), en las últimas tres décadas los estudios en el campo de la Adquisición de Segundas Lenguas (ASL) han enfocado en la enseñanza de la gramática (Whitley, 2004). La probable razón de la escasez de investigaciones dedicadas específicamente a los errores léxicos se debe a que la metodología del Análisis de Errores priorizaba los errores gramaticales y a que el interés en las cuestiones lexicales y en el proceso de adquisición del vocabulario surge a partir de los años ochenta, en un momento en el que el AE como metodología de investigación en ASL experimentaba un descenso, y se practicaba cada vez menos. Además, el hecho de que la sintaxis ofrece una idea de sistematicidad y estabilidad, mientras que los errores léxicos son muy heterogéneos, lo que dificulta el proceso de análisis, puede haber contribuido para la escasez de este tipo de investigación hasta entonces.

Pese a lo dicho anteriormente, hay investigadores que se han dedicado a la adquisición del léxico (ej.: Meara, 1984; Laufer, 1990; Agustín Llach, 2007), y específicamente en el ámbito de la adquisición del español como lengua extranjera (ELE) encontramos trabajos como los de Baralo (1997), Whitley (*op. cit.*), Higuera (2006) y Ainciburu (2007). Además, dentro de este ámbito, hay varios estudios que recurren al Análisis de Errores como metodología de investigación con el objetivo de identificar las características comunes entre los diferentes errores léxicos y poder luego generalizar (ej.: Fernández, 1997; Santos Gargallo, 1993; Alexopoulou, 2005; Palapanidi, 2011).

Como se ha visto en el apartado 2.2. (La transferencia lingüística y el léxico), en la literatura se suele distinguir entre palabras léxicas, un grupo compuesto de las palabras de clase

abierta (sustantivos, adjetivos, verbos y adverbios) y las palabras gramaticales, consideradas de clase cerrada (preposiciones, los artículos y las conjunciones). Además, los errores léxicos se conocen como aquellos que afectan a las palabras léxicas, mientras que los errores que afectan a las palabras gramaticales se examinan como errores gramaticales (Agustín Llach, 2011a).

Los estudios que se han dedicado al nivel léxico-semántico desde la perspectiva de interlengua se limitan a tres líneas de investigación (Whitley, 2004): a) la primera se dedica a estudiar cómo los alumnos infieren los significados de palabras que desconocen a través del contexto; b) la segunda se dedica a estudiar el reconocimiento de los derivados de una unidad léxica conocida, y c) la tercera estudia la producción del discurso libre oral o escrito.

Como se ha visto en el apartado 3.1. (La identificación, descripción y clasificación del error), no hay todavía un criterio uniforme de clasificación de los errores, pero los estudios sobre el léxico acostumbran a fijarse en el tipo de error, es decir, si los errores afectan a la forma o al significado y en el origen del error, es decir, si se deben a un factor interlingüístico o intralingüístico. Pero lo más importante, como destaca Agustín Llach (2011b), es que “esta doble distinción viene a ser determinante, puesto que contribuye a ofrecer al investigador una mejor visión de la organización del lexicón mental de los aprendices”.

En los análisis de los errores léxicos, se ha observado la existencia de dos posturas distintas con relación a los errores ortográficos. Hay autores que consideran los errores ortográficos como una actividad mecánica dado que se trata de un sistema arbitrario de reglas (ej.: Hamp-Lyons, 1990; Wolfe Quintero *et al.*, 1998; Martínez Arbeláiz, 2004), mientras que otros los consideran como un tipo de error léxico (Fernández, 1997; Celaya y Torras, 2001; Agustín Llach, 2011a). Este último grupo de investigadores concibe la adquisición del léxico como una actividad compleja que exige relacionar las formas con sus significados y saber usarlas en diferentes contextos discursivos respetando el registro específico (Alexopoulou, 2011). Así, entendemos no se puede disociar la competencia ortográfica de la competencia léxica.

Agustín Llach (2011a) destaca la importancia de una definición clara sobre el término error léxico, dado que para identificar un error léxico, considerado como una desviación de la norma de la L2 en la producción del alumno respecto al uso en la producción y recepción de



ítems léxicos, es necesario tener en cuenta el contexto situacional y comunicativo en el que se ha detectado el error. Además, la autora señala la importancia de una definición de lo que se está estudiando bajo la nomenclatura de errores léxicos y destaca dos categorías observadas en las investigaciones: una que reúne todos los errores léxicos en un grupo, sin diferenciarlos entre sí, y otra que se compone de subcategorías según diferentes criterios que pueden ser descriptivos, etiológicos, semánticos, epistémicos, o psicolingüísticos u orientado al proceso. Consecuentemente, la autora define error léxico (*lexical error*) como

*(...) a deviation in form and/or meaning of a target-language lexical word. Form deviations include orthographic or phonological deviations within the limits of single words, and also ignorance or syntactical restrictions which result in false collocations, for example. Meaning deviations appear when lexical items are used in contexts where they are attributed another meaning or where they violate semantic restrictions. (Agustín Llach, 2011a: 75)*

En lo que se refiere a la percepción de los estudiantes, investigadores como Grauberg (1971) y Meara (1984) ratifican en sus estudios la percepción de los estudiantes de que los errores léxicos son errores más frecuentes que los demás. Asimismo, las investigaciones que examinan los errores léxicos suelen coincidir en que éstos son los errores que más interrumpen la comunicación. De acuerdo con Santos Gargallo (1999), los lectores generalmente son capaces de juzgar contenido y lenguaje separadamente, excepto en el caso de los errores léxicos porque es en el léxico que está el contenido del mensaje y el uso de una palabra equivocada dificulta la comprensión del significado del texto y resulta en valoración negativa por parte del lector. De ahí que los errores léxicos de la interlengua de los estudiantes sean comúnmente juzgados como los más graves por los hablantes nativos, los evaluadores no-nativos y los estudiantes de L2, especialmente en lo que se refiere a la morfología (Hughes y Lascaratou, 1982). Asimismo, les parece más fácil tratar con los errores producidos que con las omisiones o elección equivocada (Ellis, 2008).

Finalmente, desde el punto de vista de la reparación y de la actuación del profesor, Olsson (1973) establece algunos criterios de prioridad para la corrección de los errores. La autora indica que, a partir de la selección de la forma equivocada, el profesor averiguaría si el error afecta la transmisión del mensaje, es decir, si interrumpe la comunicación. Sí es así, el error tendría prioridad máxima frente a otro que no afectase el significado, pero produjera irritación a su interlocutor.

Como se ha visto en este apartado, pese a la escasez de investigaciones sobre los errores léxicos en comparación con las que se han dedicado a los errores morfosintácticos, hay una variedad de estudios que han enfocado en ese aspecto lingüístico bajo diferentes perspectivas.

#### 4. LA FOSILIZACIÓN Y LA ESTABILIZACIÓN

Cuando se estudia el error en el aprendizaje de segundas lenguas (L2) es casi inevitable la referencia al fenómeno conocido como fosilización. Aunque esta tesis no se centre primordialmente en el estudio de la fosilización, sino en explorar la presencia de algunos errores que persisten a lo largo del período de formación, éste es un tema recurrente en el estudio de los errores léxicos en el ámbito de lenguas afines, por lo que se ha estimado pertinente explicar sus características y comentar algunas consideraciones que se deben tener en cuenta al realizarse investigaciones que analizan los errores.

El término fosilización ha sido introducido en el campo de la Adquisición de Segundas Lenguas (ASL) por Selinker (1972) al referirse a uno de los procesos que caracterizan la interlengua de los estudiantes de una L2, en un intento de señalar la percepción de que una gran parte de los estudiantes no llegan a tener una competencia como la de un nativo en su intento de aproximarse a la lengua meta. Según Han y Selinker (2005), esta percepción se acepta como un hecho, aunque no se reconoce la idea de que el aprendizaje de una lengua finalice totalmente, porque, para esos autores, no se puede determinar un fin en el *continuum* que representa el aprendizaje de una segunda lengua. Es por ello que Long (2003) propone el término 'estabilización' para referirse a un estadio de desarrollo de la L2 en el que se verifica la persistencia de algunas formas distintas a las de la lengua meta, cuya presencia puede ser transitoria y no constituir necesariamente una condición permanente.

Para Ellis (2008), la importancia de seguir investigando para conocer mejor los procesos de adquisición y aprendizaje reside en el interés de conocer si la adquisición de una L2 es el producto de las capacidades internas del estudiante que se modifican después de la adquisición de la L1, o si está condicionada a factores como la instrucción y la motivación del estudiante. Sin embargo, como bien destaca Sánchez Iglesias (2003), son muchos los factores que pueden originar la fosilización como son la edad, la falta del deseo de aculturación, la presión comunicativa, la falta de ocasiones de aprendizaje o el tipo de instrucción y retroalimentación (*feedback*) recibidos.

Hay un gran interés en conocer más sobre las implicaciones que tiene la edad en la adquisición de una L2. En este sentido, Ellis (*op. cit.*) afirma que aunque es más común que

los niños adquieran un acento y un dominio de la gramática como los de un nativo y no lleguen a presentar signos de fosilización, las investigaciones han revelado algunos casos en los que los estudiantes son adultos y llegaron a tener un dominio de la lengua meta similar al de un hablante nativo. Es importante destacar que los niños superan a los adultos siempre y cuando hayan tenido suficiente exposición a la lengua meta. Asimismo, el autor señala que todavía no hay un acuerdo con respecto a la teoría del período crítico, o sea, la franja de edad del niño en la que debe realizar el aprendizaje de L2 o si este período crítico puede presentar alguna variación según el aspecto de la lengua que está implicado, como la pronunciación o la gramática, por ejemplo. Además, el autor también destaca que las investigaciones no señalan una discontinuidad en el aprendizaje como resultado de la edad y hay cada vez más evidencia de que los estudiantes que empiezan a estudiar una L2 cuando adultos pueden llegar a tener una competencia como la de un nativo, lo que indica que es posible evitar la fosilización .

Sin embargo, actualmente se entiende que el objetivo principal del aprendizaje de idiomas es la comunicación, por lo que se promueve una enseñanza de lenguas basada en la interacción en situaciones comunicativas. Así, aunque el hablante de una L2 no demuestre una competencia lingüística como la de un hablante nativo, él estará capacitado para desenvolverse e interactuar de manera eficaz en una diversidad de situaciones comunicativas.

Con todo, la fosilización es un fenómeno ampliamente observado que sigue despertando el interés, no sólo de los investigadores en ASL y profesores de L2, sino también de los propios estudiantes, por estar directamente vinculada con el resultado del proceso de enseñanza-aprendizaje. Sin embargo, todavía no hay una definición consolidada de fosilización (Han, 2003; Long, 2003; Ellis, 2008) y una posible razón de ello se debe al hecho de que la investigación de este fenómeno está condicionada a limitaciones metodológicas que se comentan a continuación.

Para evidenciar que el sistema lingüístico del estudiante de L2 está estabilizado de forma permanente y demostrar que se da la fosilización de las mismas formas no correspondientes a la lengua meta después de varios años de exposición a la misma es necesario seguir algunos criterios metodológicos que se han establecido como esenciales al

realizarse investigación que pretenda demostrar la existencia de fosilización o alguna forma de cese del aprendizaje (Long, 2003; Han y Selinker, 2005): a) realizar estudios empíricos longitudinales; b) ofrecer abundantes pruebas positivas para el estudiante; c) el estudiante debe mostrar una motivación muy alta, y d) el estudiante debe tener oportunidades considerables para practicar.

Un ejemplo de un estudio longitudinal muy conocido entre los investigadores es el de Lardiere (1998) sobre Patty, una hablante de chino (L1) que vive en Canadá y que a pesar de tener una buena comprensión del inglés, su expresión oral seguía presentando errores de flexión. Aunque Patty estaba expuesta a la lengua meta en un contexto natural, los errores en su expresión oral seguían resistiendo a las correcciones y al *input* lingüístico.

De acuerdo con varios autores (Selinker y Lamendella, 1978; Selinker, 1992; Nakuma, 1998; Fidler, 2006) la fosilización es un proceso idiosincrásico, es decir, lo que se fosiliza en la interlengua de un estudiante puede no darse de igual manera en otro. Aunque se recomiende la realización de estudios longitudinales en el estudio de la estabilización que lleva a fosilización para investigar las características individuales de la interlengua de los estudiantes, hay muy pocas muestras de estudios de ese tipo en la literatura principalmente por causa de la cantidad de tiempo y energía que demandan. Por eso, muchas veces los investigadores recurren a diseños pseudolongitudinales.

Es debido a las limitaciones metodológicas que afectan al desarrollo de estudios longitudinales que se proponen a evidenciar la fosilización que Long (*op.cit.*) afirma que es casi imposible demostrarla. Por eso, en lugar de estudiar la fosilización, el autor propone el estudio de este fenómeno bajo el término estabilización, ya que no constituye una condición permanente, sino un estadio del desarrollo de la L2 en el que la oscilación se ha detenido temporalmente. Así, es viable averiguar si algunas formas distintas a las de la lengua meta persisten a lo largo de un período determinado y demostrar la estabilización empíricamente. Además, como bien señala Ellis (2008), el término ‘estabilización’ no implica la idea de fracaso presente en el término ‘fosilización’. Sin embargo, pese a todo lo dicho, ‘fosilización’ es el término más difundido en la literatura, y se suele utilizar para referirse también a la ‘estabilización’ del sistema lingüístico del estudiante de L2.

Han (2004) distingue tres tipos de investigación que se han dedicado al estudio de la fosilización: a) las que han buscado formas no correspondientes a la L2 en la interlengua de los estudiantes; b) las que han focalizado la atención en los errores más comunes entre estudiantes que comparten la misma L1, y c) las que han estudiado los errores encontrados en la producción de estudiantes de niveles avanzados en un intento de explicar que los errores que se encuentran en ese nivel de competencia deberían ser los errores más persistentes y con más propensión a la fosilización. Así, los investigadores han seguido básicamente los siguientes enfoques metodológicos (Han, *op. cit.*: 16): a) longitudinal; b) error común; c) estudiante avanzado; d) *feedback* correctivo, y e) tiempo de residencia.

Dado que el enfoque metodológico pseudolongitudinal ha sido uno de los más utilizados porque permite al investigador analizar muestras de la interlengua de los estudiantes sin la necesidad de acompañarlos a lo largo de todo su período de formación, por ejemplo, comentamos sucintamente algunas de las ventajas y desventajas que la utilización de este enfoque aporta al estudio de la fosilización. El propósito de señalar las desventajas del enfoque pseudolongitudinal no es desaconsejar su uso, sino más bien advertir al investigador de las particularidades propias de dicho enfoque que se deben tener en cuenta al emplearlo.

Un ejemplo de ello es una investigación pseudolongitudinal sobre el error común que realizó Kellerman (1989). En su estudio, el autor reivindica que los datos obtenidos con estudiantes de una comunidad específica en diferentes niveles de competencia revelan evidencia de que determinadas estructuras lingüísticas tienen más propensión a la interferencia de la L1. Sin embargo, según Han (2004) la evidencia obtenida en estudios pseudolongitudinales sólo ofrece una visión general de los hechos, es decir, no se sabe qué ocurre en la interlengua de cada uno de los participantes del estudio. Eso disminuye la credibilidad sobre el error común como índice de fosilización en la interlengua de cada estudiante de L2 o indica erróneamente que el error común se va a fosilizar en todos los estudiantes que comparten la misma interlengua.

Otro grupo de investigadores (ej.: Hyltenstam, 1988) se han centrado en analizar los errores que todavía permanecían en la interlengua de estudiantes avanzados o con una competencia próxima a la de un nativo. Su idea era la de que esos errores persistían tras un

largo período de estabilización. Sin embargo, Han (*op. cit.*) llama la atención sobre el hecho de que si los estudiantes han sido capaces de aproximar tanto su interlengua a la L2, a lo mejor son capaces de hacerlo todavía más. Además, el autor afirma que sería necesario tener evidencia longitudinal para asegurarnos de que las formas desviadas son realmente producto de una larga estabilización y no una reestructuración reciente.

A partir de esas consideraciones iniciales sobre la investigación del fenómeno de la fosilización, pasamos a comentar algunas de las clasificaciones que se han propuesto hasta el momento.

#### 4.1. Las categorías de la fosilización

La fosilización puede ser ‘global’ cuando afecta a toda la interlengua, o ‘local’ si se observa sólo en algunos aspectos de la lengua, como por ejemplo en las formas de pasado (Han, 2004). Es interesante señalar que el desarrollo sintáctico y el gramatical no avanzan de igual manera durante el proceso de adquisición de una L2, por lo que es posible que mientras se verifique estabilización de algunos elementos, es posible que otros todavía estén en desarrollo (Lardiere, 1998).

La fosilización puede también referirse tanto a un ‘producto’, que puede afectar a la interlengua tanto de manera local como global, como a un ‘proceso’ en el que no se habla de un cese en el desarrollo de un determinado aspecto lingüístico, sino más bien en el que se verifica una posible tendencia al cese del desarrollo de la interlengua (Fidler, 2006).

Han y Selinker (2005) categorizan la fosilización en dos niveles, el ‘cognitivo’ y el ‘empírico’. En el nivel cognitivo, la fosilización se refiere a los procesos cognitivos, o a los mecanismos subyacentes que producen permanentemente formas estabilizadas de interlengua. En el nivel empírico, la fosilización se refiere a las formas estabilizadas de interlengua que permanecen a lo largo del tiempo en el discurso oral o escrito, sin que influya en ello el *input* o lo que haga el estudiante.

Siguiendo la clasificación presentada inicialmente por Selinker y Lamendella (1978), Vázquez (2010) distingue los errores en dos tipos: a) aquéllos en los que la fosilización puede ser ‘colectiva’, cuando se observan errores comunes en un grupo de estudiantes que suelen

compartir la misma L1, o b) 'idiosincrásica', cuando los errores identificados están presentes solamente en la producción de algunos individuos. Según la autora, los errores individuales revelan la internalización de las reglas de la L2 y desaparecen por sí mismos, mientras que los errores colectivos presentan mayor resistencia a desaparecer. Además, los errores colectivos pueden agruparse en fosilizados, cuando su aparición es inesperada y se refieren a reglas aprendidas y utilizadas correctamente en otros casos, y transitorios, aquéllos que suelen desaparecer con el tiempo, pero son característicos de determinadas etapas del aprendizaje (Vázquez, *op. cit.*: 164-165).

#### 4.2. La fosilización y la influencia de la L1

De acuerdo con Han y Selinker (2005), en los estudios en los que se analizan los errores que son propios de estudiantes que comparten la misma L1 a lo largo de diferentes niveles de competencia hay dos supuestos principales: a) los errores que tipifican toda una comunidad de estudiantes de L2 que comparten una misma L1 son los candidatos más fuertes para fosilización, y b) los errores que no son comunes a la comunidad, pero también permanecen en los miembros que tienen una competencia más avanzada, son indicativos de fosilización.

Además, estos autores también destacan que en el caso de lenguas afines, es decir, cuando la L1 del estudiante y la lengua meta (L2) son lenguas tipológicamente similares, se verifica la presencia de elementos fosilizados en la producción de los estudiantes que ya alcanzaron un nivel intermedio/avanzado de competencia. Esto puede ser así porque el estudiante no es capaz de percibir que en la lengua meta hay otra palabra u otra estructura que corresponde a lo que quiere decir, o bien porque el estudiante es capaz de desenvolverse de manera adecuada en la L2 desde el punto de vista comunicativo y no tiene la oportunidad de percibir que su interlengua todavía tiene elementos de la L1, ya que sus interlocutores pueden comprenderle y no se observa interrupción en el proceso comunicativo.



## 5. LA PRODUCCIÓN ESCRITA Y LAS MEDIDAS DE CALIDAD

En este capítulo se comentan brevemente las características de los textos escritos en el contexto de aprendizaje de segundas lenguas (L2), una vez que la presente investigación se centra en el estudio de los errores en las palabras léxicas en textos escritos. Además, se hace un repaso de las principales medidas de análisis de la calidad del texto escrito: la fluidez, la diversidad léxica y la corrección.

### 5.1. La producción escrita en el ámbito de las segundas lenguas

Actualmente, la producción escrita sigue teniendo importancia en el enfoque comunicativo, ya no como forma de control de la expresión oral, sino como una de las formas de expresión que forman parte del proceso de comunicación.

De acuerdo con Alba Quiñones (2009), de las cuatro destrezas comunicativas (expresión oral y escrita y comprensión oral y escrita), los estudios han focalizado principalmente en las expresivas. Según la autora, esto se debe, fundamentalmente, a que se puede cuantificar más fácilmente la producción que la recepción. Asimismo, los estudios normalmente se concentran en el campo de la producción escrita (espontánea o dirigida, formal o informal) obtenida a través de distintos procedimientos como pruebas, cuestionarios, composiciones libres o dirigidas, cartas, e incluso el uso de las pruebas de los Diplomas de Español del Instituto Cervantes como se puede ver en la tesis de Sánchez Iglesias (2003).

Cornaire y Raymond (1999) caracterizan los textos escritos en segundas lenguas (L2) como textos muy cortos, que tienen un vocabulario reducido, una sintaxis simple y que presentan más errores que los textos escritos en lengua materna (L1). Aún según estas mismas autoras, los procesos de escritura en L2 comparten características como un repertorio de estrategias limitado o inadecuado, una competencia lingüística limitada y la necesidad de un tiempo de redacción más largo. Esta competencia lingüística limitada se refleja, muchas veces, en los errores.

Al comparar la ocurrencia de los errores en textos escritos y en textos orales, vemos que en lo que se refiere a la gravedad del error, se suelen considerar como más graves los errores producidos en textos escritos que los errores que ocurren en los discursos orales. Esto es así

debido al carácter efímero e inmediato de la oralidad, lo que hace que la gente los tolere más. Asimismo, en cuanto al tema de la percepción de los errores, James (1998) los clasifica en dos tipos: errores explícitos (*overt errors*) y errores encubiertos (*covert errors*) e indica que los errores explícitos escritos son más fáciles de percibir, no sólo porque lo que está escrito tiene un carácter permanente mientras que lo hablado es pasajero, pero también porque hay más expectativas sobre el desempeño de lo escrito.

En relación con la utilización del texto escrito como instrumento de investigación, Kellerman (2001) propone que los estudios de transferencia se hagan a través de las narrativas, dado que constituyen un ejemplo particularmente rico de datos y que ofrecen un contexto en el que se pueden examinar los errores tanto bajo el aspecto lingüístico como conceptual. Además, Agustín Llach (2011a) destaca que el proceso de desarrollo de la escritura tiene un carácter multidisciplinario y el escribir se puede considerar bajo perspectivas variadas, es decir, como producto o proceso, como una actividad social o cognitiva y como una actividad orientada al contenido o a la forma.

Además, las actividades de escritura han sido desde siempre una actividad que forma parte del aprendizaje de lenguas en contextos formales de enseñanza (Navés *et al.*, 2005). Así, la utilización de pruebas escritas como herramienta de obtención de datos no representa una situación de estrés para el alumno, ya que se trata de una actividad a la que está habituado a realizar.

## 5.2. La fluidez

La palabra 'fluidez' ha sido utilizada en la literatura para describir diferentes fenómenos. Un ejemplo de ello es la utilización del término fluidez para representar la velocidad con la que una persona habla o escribe, o el grado de coherencia, complejidad y creatividad en un discurso o en una producción escrita. Sin embargo, como observa Wolfe-Quintero *et al.* (1998), la dificultad asociada a la idea de fluidez se puede solucionar cuando se analiza fluidez independientemente de diversidad y corrección.

En el ámbito de la enseñanza de segundas lenguas (ASL), se ha observado que a medida que ganan competencia en la L2, los estudiantes son capaces de escribir con más fluidez, es

decir, utilizan más palabras en su producción oral o escrita, en el mismo período de tiempo (Wolfe-Quintero *et al.*, 1998; Agustín Llach, 2011a). Así, cuanto mayor una producción escrita, mayor el grado de fluidez (Jarvis *et al.*, 2003). Consecuentemente, la fluidez está directamente relacionada con la cantidad de palabras escritas.

Sin embargo, es importante destacar que la fluidez léxica o el tamaño de la producción escrita no es una medida directa de la calidad de lo escrito, pero se ha detectado que una mayor producción escrita está directamente asociada a una mayor riqueza de información (Reid, 1990). En este sentido, Agustín Llach (*op. cit.*) observa que el vocabulario es el indicador más significativo para medir la calidad de un texto. Además, varios estudios confirman la relación entre conocimiento de vocabulario y la calidad de lo escrito. Un ejemplo de ello es el estudio de Engber (1995), en el que la autora procuraba encontrar una relación entre la competencia léxica, medida a través de la variedad léxica y el número de errores léxicos, y la evaluación holística sobre la calidad general de los textos. Para ello, se analizaron la diversidad léxica, la diversidad léxica sin errores, la densidad léxica, los errores léxicos y los resultados se correlacionaron con la evaluación holística de los textos escritos.

De acuerdo con Wolfe-Quintero *et al.* (*op.cit.*), una mayor fluidez léxica significa una producción con más palabras y más estructuras en un limitado periodo de tiempo, mientras que una baja fluidez léxica se observa cuando hay pocas palabras y estructuras en un texto. Luego, fluidez léxica no es una medida de sofisticación o precisión de las palabras o formación de las estructuras, pero es la medida del número de palabras o estructuras escritas en un determinado período de tiempo. De ahí que, la forma más sencilla para medir la fluidez léxica es contabilizar el número de frases por texto y/o el número de palabras por texto (N de palabras por texto). Con todo, en la literatura se puede encontrar una diversidad de medidas, como por ejemplo el tamaño de la frase sin error y las ratios de palabras por minuto y número de palabras por frase.

### 5.3. La diversidad léxica

La diversidad léxica puede ser definida como el rango y variedad de vocabulario presente en un texto. Por eso, la diversidad léxica también es conocida como variedad léxica. Asimismo, la definición de diversidad léxica varía en cada estudio. Mientras que un grupo de

estudiosos considera diversidad y riqueza léxica como términos sinónimos, otros consideran que la diversidad léxica es uno de los cuatro componentes de la riqueza léxica (Read, 2000). Para este último grupo de investigadores, el uso efectivo del vocabulario en la escritura consiste en variedad léxica, sofisticación léxica, densidad léxica y también en un número reducido de errores. Durán *et al.*, (2004) destacan que se han realizado estudios sobre la diversidad léxica como variable en muchas áreas de la investigación lingüística como por ejemplo en el análisis de documentos históricos, estudios de estilo del autor y determinación de fecha de trabajos literarios y lingüística forense. En cuanto al estudio de la diversidad léxica en la enseñanza de lenguas, los autores también citan el ámbito de las segundas lenguas y bilingüismo, y la evaluación del lenguaje, tanto de la lengua materna (L1), como de segundas lenguas (L2), así como en una multitud de estudios sobre discapacidades del lenguaje. Aún según estos autores, cabe destacar que pese a la dificultad matemática en establecer métodos de medición, los estudios empíricos revelaron que las variables de diversidad léxica son válidas como índices de desarrollo y también como medidas teóricamente motivadas en determinar el perfil de una amplia gama de discapacidades del lenguaje.

Así como los términos utilizados para referirse a la diversidad léxica, las medidas que se utilizan para calcularla varían mucho. Laufer (2003), por ejemplo, utiliza el término diversidad léxica para referirse a la combinación entre la variedad léxica y la sofisticación léxica. Sin embargo, Malvern *et al.* (2004), utilizan el término 'diversidad léxica' para hacer referencia a la variedad léxica, es decir, el número de palabras diferentes o únicas identificadas en cada texto (*types*). La preferencia de esos autores por el término diversidad léxica se debe al hecho de que es el más utilizado en la literatura que se dedica al estudio del lenguaje infantil, que ha sido el punto de partida de su interés en el área.

La medida más sencilla que se ha utilizado para medir la diversidad léxica es el cómputo del número de palabras diferentes (NPD), aunque esta medida haya indicado una alta correlación con la edad y también haya sido eficaz en la diferenciación entre niños con un desarrollo lingüístico estándar de los que tenían alguna discapacidad del lenguaje, se ha visto que su utilización es limitada cuando se comparan textos de tamaños diferentes (Wolfe-Quintero *et al.*, 1998; Dewaele y Pavlenko, 2003; Durán *et al.*, *op. cit.*). Se han realizado estudios con diferentes procedimientos de estandarización en un intento de poder

comparar textos de tamaños diferentes, pero la variedad de procedimientos utilizados producen resultados muy distintos lo que mina la utilización del NPD como medida fiable de diversidad léxica.

Una medida léxica muy habitual es la ratio *type/token* (RTT), que utiliza el cómputo del número de palabras diferentes o únicas (*types*) dividido por el cómputo del número total de palabras (*tokens*) de una muestra (Linnarud, 1983). Con todo, la RTT se ve afectada por el tamaño del texto, es decir, cuanto más palabras en el texto (más *tokens*), más posibilidades hay de que las palabras sean repetidas (menos *types*). Esto ocasiona una disminución en los valores de la ratio, que pueden variar entre 0 (cero) y 1 (uno). Así, los textos más largos tendrán un valor más bajo de la ratio de diversidad léxica. Eso hace que sea muy difícil comparar las medidas de diversidad léxica de una muestra con textos de tamaños diferentes (Vermeer, 2000; Jarvis, 2002). De todos modos, la interpretación del análisis de la diversidad léxica siempre va a depender de la definición del investigador sobre qué considera como palabra repetida (*token*) y como palabra diferente o única (*type*) (Malvern *et al.*, 2004).

Con el objetivo de superar estas dificultades, se han propuesto soluciones como, por ejemplo, computar sólo las 100 primeras palabras de cada texto (Biber, 1988), pero la idea de no examinar el texto como un todo ofrecía resultados parciales. Asimismo, varios autores han desarrollado fórmulas y ecuaciones matemáticas que procuraban anular el efecto del tamaño de la muestra. A modo de ejemplo, comentamos el índice de Guiraud (1960) que, en un intento de corregir el efecto del tamaño del texto en la ratio, desarrolló la RTT cuadrada (*Root TTR*), es decir, la ratio del número de palabras diferentes (*types*) dividido por la raíz cuadrada del número total de palabras (*tokens*) ( $RTTR = \frac{types}{\sqrt{tokens}}$ ). Carroll (1964) a su vez, desarrolló la medida de la RTT corregida (RTTC), que consistía en una ratio del número de *types* dividido por la raíz cuadrada del doble del número total de palabras ( $RTTC = \frac{types}{\sqrt{2*tokens}}$ ). Aunque algunos investigadores (Arnaud, 1984; Malvern *et al.*, *op. cit.*) no estén de acuerdo con la idea de que ese cambio corrija el efecto del tamaño de los textos, Jarvis (2002) defiende que la medida de la RTTC funciona bien para textos que tienen entre 1.000 y 10.000 palabras.

La creación de una herramienta informática (vocD) que incorpora una fórmula matemática especialmente desarrollada para calcular la diversidad léxica D (McKee *et al.*, 2000) ha

impulsado el desarrollo de varias investigaciones que se dedicaron a examinar la diversidad léxica (ej.: Richards y Malvern, 1999; Malvern y Richards, 2000; Jarvis, 2002; Strömquist *et al.*, 2002; Wright *et al.*, 2003). Esto se debe principalmente a la facilidad que supone la utilización de esta herramienta informática para calcular la diversidad léxica. VocD hace parte de un conjunto de herramientas que se han creado con el propósito de auxiliar en el estudio de la adquisición del lenguaje infantil y que forman parte de un programa informático conocido como CHILDES (McWhinney, 2000).

Con todo, pese a que haya evidencia de que la medida de diversidad D no se ve afectada por el tamaño de la muestra, tras realizar un estudio en el que comparaba la utilización de varias medidas de diversidad léxicas, Jarvis (2002) destaca que vocD sólo es fiable para medir la densidad léxica en textos que tienen entre 200 y 400 palabras. Así, este autor averigua que el índice Uber U (Dugast, 1978, 1979, 1980) es la medida más recomendable para examinar la diversidad léxica en textos pequeños (figura 4).

$$U = \frac{(\log \text{tokens})^2}{\log \text{tokens} - \log \text{types}}$$

Figura 4. Medida de diversidad léxica Uber U.

El índice Uber U es una transformación algebraica de la RTT y ofrece un ajuste preciso de la RTT *versus* la curva N (número total de palabras), dado que una gran diversidad léxica pasado un determinado límite de palabras, puede indicar una baja calidad del escrito (Dewaele y Pavlenko, 2003). Asimismo, Vermeer (2000) explica que este índice ofrece resultados adecuados solamente para textos de hasta 3.000 palabras.

Strömquist *et al.* (2002) señalan que los valores absolutos de los índices de medidas de diversidad léxica pueden variar mucho según la tipología lingüística de la lengua estudiada. Por otro lado, como observan Simon-Cerejido y Gutiérrez-Clellen (2009) la mejor metodología para estudiar la diversidad léxica en lengua española aún no está determinada porque, al contrario del inglés, el español presenta una gran diversidad de verbos, artículos y adjetivos flexionados que dificultan el simple cómputo de las palabras.

#### 5.4. La calidad del texto escrito

Para medir la calidad del texto escrito se suelen utilizar conjuntamente una serie de medidas léxicas como la fluidez (ej.: N total de palabras por texto), la diversidad léxica (ej.: RTT, índice Uber U) y la corrección, medida a través del cómputo del número de errores (Polio, 1997).

En su estudio sobre el papel de los errores léxicos como indicadores de la calidad del texto, Agustín Llach (2007) explora la relación entre el tipo de error léxico y la calidad general del escrito por un lado, y por otro, entre el tipo de error léxico y el ámbito del vocabulario, pertinencia y precisión. Según la autora, tanto los errores de manera general, como los errores léxicos en particular, son un criterio importante en la medición de la calidad de lo escrito y que sirven como indicadores para predecir la nota del texto, aunque también hay otros factores que tienen influencia en la nota final. Asimismo, la autora afirma que el vocabulario también es un criterio muy utilizado para la medición de la calidad de los textos escritos. Además, la autora afirma que en el contexto pedagógico se observa un amplio uso de los errores léxicos porque ellos son elementos útiles como indicadores de la calidad del escrito de los estudiantes, del progreso léxico, de la competencia léxica de los estudiantes y de su resultado académico. Un ejemplo de ello es la diversidad léxica que ha sido considerada una medida fiable, dado que la competencia léxica es un indicador de la calidad de la composición (Santos, 1988).

Una investigación muy relevante en el ámbito de la ASL que analiza la relación entre la competencia léxica y la calidad del texto escrito es la que realizó Engber (1995). Como hemos mencionado en el apartado referente a la fluidez, esta investigadora analizó redacciones de estudiantes con nivel intermedio a intermedio alto y encontró resultados que indicaban que la diversidad léxica en conjunto con la corrección léxica tienen un efecto significativo en el lector con respecto a la calidad de los textos. Esto es así porque se verificó una correlación existente entre la puntuación de cada redacción y los resultados de la medida de variedad léxica libre de errores. Por lo tanto, se ha visto que la producción escrita de los sujetos más competentes en L2 presenta mayor diversidad léxica y menor número de errores. La autora también sugiere que la variedad léxica puede ser una estrategia propia de los niveles intermedios para expresar contenido, mientras que en los niveles avanzados los

estudiantes tienen mayor grado de precisión léxica, lo que explicaría índices bajos de diversidad.

También de acuerdo con Jarvis *et al.* (2003), la diversidad léxica asociada a la fluidez son características presentes en los textos evaluados como buenos. Asimismo, en un estudio más reciente, Mellor (2011) encontró una gran correlación entre los resultados de la evaluación de la calidad de textos escritos en inglés (L2) con el análisis de la cantidad (fluidez) y contenido (diversidad léxica) de las redacciones.



## PARTE II: EL CONTEXTO GENERAL DE LA INVESTIGACIÓN

### 6. CONTEXTO LINGÜÍSTICO

Así como lo expresa Soto (2002), pensamos que no se puede dejar de tener en cuenta las variaciones propias de cada uno de los sistemas lingüísticos en contacto cuando tratamos del tema de la transferencia. Por tanto, nos parece pertinente tejer algunos comentarios con respecto a la L1 de los sujetos participantes de esta investigación, la variedad brasileña de la lengua portuguesa, para dejar constancia de su riqueza y diversidad.

#### 6.1. Las variedades del portugués

En el mundo académico brasileño cada vez más se defiende la idea de que el portugués de Brasil y el de Portugal constituyen dos lenguas distintas (Celada y Rodrigues, 2005). Por otro lado, se observa un esfuerzo de las Comunidades de los Países de Lengua Portuguesa (CPLP) de unificar la lengua portuguesa a fin de facilitar su divulgación y la participación en eventos internacionales a través de acuerdos ortográficos como el que se firmó en 1990 y entró en vigor en el 2009 (Moura Neves, 2010).

La lengua portuguesa se habla en varios países<sup>19</sup>, entre ellos Brasil, que abarca un territorio muy amplio en el que conviven diversas etnias, lo que conlleva a la existencia de variedades sociolingüísticas y regionales. La marca de la variedad regional se observa principalmente en el acento y en el uso de vocabulario específico. Estas características lingüísticas propias de cada región pueden influir en el aprendizaje de una segunda lengua (L2) y la interlengua de los estudiantes podrían presentar características distintas si se investiga la producción de participantes de diferentes partes del país.

En su estudio sobre las dificultades de pronunciación que presentan, de manera general, los estudiantes brasileños de español como lengua extranjera (ELE), Barbosa (2008) observa que en los niveles iniciales de la enseñanza, algunos estudiantes presentan problemas en la

---

<sup>19</sup> La lengua portuguesa se habla en América (Brasil), Europa (Portugal), África (Angola, Cabo Verde, Guinea-Bissau, Guinea Ecuatorial, Mozambique, Santo Tomé y Príncipe) y en Asia (Timor Oriental).

producción de fonemas nasales, en la apertura y cierre de fonemas vocálicos, en los encuentros vocálicos en diptongos crecientes, en algunos fonemas y alófonos oclusivos y fricativos, así como dificultades en la realización de la vibrante múltiple y los alófonos vibrantes en general.

Sin embargo, hay algunos rasgos de la pronunciación del portugués variedad brasileña que no se observan de manera uniforme en todo el país. Es el caso, por ejemplo, de la tendencia en transformar el fonema /t/, fonema oclusivo, dental, sordo en portugués y español, en [tʃ], alófono portugués palatal africado sordo, cuando va seguido del fonema /i/ (ej.: discutir) (Masip, 1999). La producción de este alófono no se verifica, en general, en los rasgos fonológicos de los habitantes de la región Nordeste de Brasil, ni específicamente en los hablantes del estado de Pernambuco, lugar de origen de los participantes de esta investigación. Es importante destacar que, aunque el fonema /t/ exista en portugués, los hablantes de las regiones que lo transforman en el alófono [tʃ] presentan rasgos de interferencia en su producción oral en español. Esta interferencia en la producción oral causa, en muchos casos, una dificultad de comprensión en situaciones de comunicación oral con hablantes nativos de español.

Asimismo, en Brasil se verifica la existencia de diferencias lingüísticas con relación a la lengua escrita y la lengua oral que pueden interferir en la interlengua de los estudiantes. Varios investigadores se han dedicado al estudio del portugués oral en Brasil, e incluso hay gramáticas específicas sobre el *portugués falado*<sup>20</sup> (Castilho, 1989). En la lengua oral no se siguen muchas de las reglas gramaticales que se enseñan como norma culta. Esta ambigüedad de reglas lingüísticas según el registro, oral o escrito, afecta el aprendizaje de una L2, principalmente de una lengua próxima como es el caso del español por lusohablantes brasileños. Esto es así porque el estudiante recurre a su L1 durante su proceso de aprendizaje, principalmente en los niveles iniciales, cuando está probando y haciendo hipótesis sobre las reglas de la lengua que está aprendiendo. A modo de ejemplo, destacamos la colocación pronominal. La norma culta de la lengua portuguesa indica que no se puede empezar una oración con pronombre átono. Sin embargo, este uso de la lengua se restringe muchas veces al ámbito escrito. En la lengua oral, los hablantes brasileños suelen

---

<sup>20</sup> El libro es el resultado de los trabajos del “Projeto Gramática do Português Falado” que tiene como objetivo elaborar una gramática de referencia de la variedad culta del portugués oral en Brasil.

anteponer el pronombre al verbo con frecuencia, aunque éste se encuentre al principio de la frase y en lugar de utilizar la norma culta (ej.: *Diga-me./Vou-me embora.*) es muy común el empleo de la forma no culta (ej.: *Me\* diga.*) o incluso la supresión del pronombre (ej.: *Vou\* embora.*). En el caso de la colocación pronominal, el estudiante duda al principio de su proceso de aprendizaje del español a qué regla debe recurrir, si a la norma culta de su L1, utilizada principalmente en el registro escrito, o a la norma utilizada más frecuentemente en el registro oral, que sufre el estigma de norma no culta. Esto se agrava principalmente cuando el alumno empieza a tener contacto con formas y reglas de la L2 que contradicen la norma culta de su L1 (ej.: *Me voy.*), una vez que la norma de la lengua española indica que los pronombres deben ir antepuestos al verbo y sólo se posponen al mismo si éste se encuentra en sus formas de imperativo afirmativo (ej.: *dígame*), gerundio (ej.: *diciéndome*) o infinitivo (ej.: *decirme*), independientemente de su posición en la frase. Como se ha dicho anteriormente, esta norma gramatical del español (L2) genera un conflicto con las reglas que tiene el estudiante de su propia lengua, el portugués, una vez que la norma culta de la L2 contraría la norma culta de su L1 (Milani, 1999).

No se ha profundizado más en las características de la lengua hablada porque el objetivo principal de este trabajo es estudiar los errores en las palabras léxicas en un *corpus* de lengua escrita y porque los participantes del estudio son alumnos universitarios, lo que implica que tienen un buen dominio de la norma culta de su L1. Sin embargo, nos ha parecido importante mencionar este rasgo propio de la variedad lingüística de los participantes que comprenden la muestra de esta investigación, dado que al tratarse del aprendizaje de una lengua próxima, es importante tener presentes también las particularidades de la L1 en el análisis de la interlengua.

En el siguiente apartado comentamos sobre la presencia de la lengua española y su enseñanza en Brasil.



## 7. LA ENSEÑANZA DE ESPAÑOL EN BRASIL

Pese a la presencia significativa de españoles en Brasil y de su situación de vecindad con países hispanohablantes, tanto el uso como el aprendizaje del español en el país han sido muy reducidos. Moreno Fernández (2005) hace un recorrido sobre la situación histórica del español en Brasil y constata que, durante muchos años, el español casi no había figurado en el sistema educativo brasileño.

Sin embargo, desde finales de los años noventa esta situación ha cambiado. Se nota cada vez más el interés por la lengua española en el país. Su creciente importancia, impulsada en principio por el Mercosur<sup>21</sup>, viene determinando su inclusión en los currículos escolares y aumentando el fortalecimiento de las relaciones entre sus habitantes en los ámbitos cultural, social, económico y político. Celada y Rodrigues (2005: 1), destacan que “la firma e implementación del Mercosur tuvieron un impacto interesante en la historia marcada en Brasil por una cierta inercia con relación al español, lengua vista como *parecida, fácil o de estudio innecesario*”.

La llegada del Instituto Cervantes a varias capitales del país ha dado un gran impulso al aprendizaje del español en la enseñanza no reglada. En 1998 se inauguró el primer centro en la ciudad de São Paulo y en 2001 el segundo en Río de Janeiro. En un primer momento, ambas sedes impartían cursos de formación permanente, metodología y didáctica para profesores de español. En un segundo momento, comenzaron a ofrecer cursos regulares y especiales de español como lengua extranjera, además de un amplio programa de actividades culturales (Basterrechea Moreno, 2007). Posteriormente, diez años después, se inauguraron los centros de Recife, Salvador de Bahía, Brasilia, Curitiba, Porto Alegre, Florianópolis y Belo Horizonte, que sustituyeron los antiguos *Centros Culturais Brasil-Espanha* de la Agencia Española de Cooperación Internacional (AECI) (Pérez de Obanos, 2010).

Lo cierto es que se verificó un cambio respecto al interés de los brasileños por el español, porque pasaron a considerarlo (además del inglés) como lengua vehicular en los negocios,

---

<sup>21</sup> Tratado de Asunción, firmado por Brasil, Argentina, Paraguay y Uruguay el 26 de marzo de 1991 que dio origen al Mercado Común del Sur – MERCOSUR. Tiene como idiomas oficiales el español y el portugués. Brasil representa el 70% del PIB de la región.

en el trabajo y como herramienta de acceso a Europa y a los Estados Unidos (Celada y Rodrigues, *op. cit.*). El español es una de las lenguas extranjeras que más generan interés en el país, lo que se refleja en la demanda de cursos en todos los niveles (Moreno Fernández, *op. cit.*) y como podemos observar en la creciente opción por esta lengua en los concursos *vestibular*<sup>22</sup> de las universidades brasileñas.

Tanto el Mercosur como la inclusión del español en el *vestibular* son factores que contribuyeron contundentemente para el crecimiento de la enseñanza de la lengua española en Brasil y consecuentemente para la aprobación del Proyecto de Ley 3.987/2000 de autoría del Diputado Átila Lira en el Congreso Nacional, y su posterior sanción el 5 de agosto de 2005 por el Presidente de la República Luiz Inácio Lula da Silva. Esta ley hace obligatoria la oferta de la lengua española en las escuelas públicas y privadas de la enseñanza media de todo el país desde el 2010. El artículo primero instituye la oferta obligatoria de la materia lengua española en el bachillerato (*ensino médio*), sin embargo, para los alumnos, el estudio de esta materia es facultativo. En los currículos de los últimos cuatro años de la enseñanza secundaria, la inclusión del español será optativa, es decir, la oferta del curso de español es obligatoria, pero el alumno puede elegir entre estudiar inglés o español.

El informe *Datos y Cifras* de la Consejería de Educación de la Embajada de España (Morillo, 1998) refleja una deficiencia de profesionales especialmente formados y dedicados a la enseñanza del español. Se espera que la nueva ley, conocida como ley del español o ley 11.161 de 2005, represente nuevas perspectivas de trabajo para los alumnos de los cursos de Letras que se dedican al estudio de la lengua española, porque plantea el crecimiento de ofertas para docentes de este idioma. Sin embargo, la falta de medios técnicos, de apoyo bibliográfico, de material auténtico tanto impreso como audiovisual, y, principalmente, de profesorado cualificado representa una dificultad al cumplimiento de la nueva ley. Cinco años después de la sanción de la ley, “según el *Instituto Nacional de Estudos e Pesquisas Educacionais* (órgano perteneciente al Ministerio de Educación Brasileño), el 75% de las 25.700 escuelas públicas y privadas brasileñas no está en condiciones de cumplir la ley” (Pérez de Obanos, 2010: 58).

---

<sup>22</sup> Examen de admisión a la universidad en Brasil.

El principal motivo que explica esta situación sigue siendo la deficiencia de profesores licenciados en lengua española. Para la implementación de la Ley del Español en las escuelas públicas y privadas, hacen falta más de 29.000 profesores (Candeas, 2007). Los datos del Ministerio de la Educación (MEC) muestran que hay en Brasil 12.700 profesores de lengua española, y casi 8 millones de alumnos en la enseñanza secundaria; además, en el ámbito universitario, aparte del curso de Letras, se destaca la expansión de la enseñanza de la lengua española en las carreras de Turismo, de Relaciones Internacionales y de Comercio Exterior (Pérez de Obanos, 2010).

Según el anuario 'El español en el mundo' (Instituto Cervantes, 2009), la enseñanza del español atraviesa el mejor momento de su historia, tanto en calidad como en cantidad. Aún de acuerdo con este informe, la universidad consigue atender la demanda existente de profesorado de español. Sin embargo, todavía hacen falta inversiones y más medios económicos para atraer profesionales y consolidar la carrera del profesor de lengua española.

### 7.1. La enseñanza del español en Pernambuco

Tras la contextualización general que acabamos de presentar sobre la enseñanza de la lengua española en Brasil, nos centramos en describir brevemente el panorama específico de la enseñanza del español en el estado de Pernambuco. Para ello, recurrimos al informe de Martínez-Cachero Laseca (2009) sobre la enseñanza del español en el sistema educativo brasileño porque es un documento que presenta una clara descripción de la situación de la enseñanza de la lengua española que podemos corroborar a partir de nuestra propia experiencia.

Aunque el informe Datos y Cifras (1998) no presente datos sobre la enseñanza de español en el sistema educativo público en contexto no universitario de Pernambuco, el informe MEC/SEB (2005, en Martínez-Cachero Laseca, *op. cit.*: 188) indica que este estado de Brasil tenía 14 profesores de español trabajando en 13 escuelas públicas. Ese cambio representa, por lo tanto, una mejora con respecto a la implementación de la enseñanza de la lengua española en el sistema educativo tras la aprobación de la ley 11.161,

*embora não resulte fácil saber a situação exata já que, como ocorre em outros Estados, a Secretaria de Educação de Pernambuco (SEP) não possui dados sobre seus professores de espanhol. Esse fato é possível já que os professores que lecionam a disciplina são, ocasionalmente, professores de outras disciplinas, especialmente de português, e não constam nas estatísticas como professores de espanhol. Tudo isso nos impede de saber com certeza seu número, assim como quantos alunos recebem aulas de espanhol. (Martínez-Cachero Laseca, 2009: 188-189).*

Asimismo, el autor explica que en el 2005 se ofertaron 35 plazas para profesores de español con Licenciatura en Letras y especialización en lengua española. Además, en el 2006 se ofertaron tres nuevas plazas y en el 2008, siete, y finalmente, otra oposición, en el 2009 ofertó trece plazas para profesor de español.

En el ámbito de las escuelas particulares, hace muchos años se ofrecen clases de lengua española, algunas a partir de la enseñanza media y otras en el último año o en los cursos preparatorios para el *vestibular*. En algunos casos, se ofrecen clases de lengua española a partir de la enseñanza fundamental (*séries 5<sup>a</sup> a 8<sup>a</sup>*) como materia optativa o a través de cursos de lengua en academias que hacen acuerdos con las escuelas (Martínez-Cachero Laseca, 2009). En lo que se refiere a los profesores de las escuelas particulares, el autor explica que se observa una gran variedad, porque algunos tienen la licenciatura en Letras con especialización en español, pero hay también profesores que dan clases porque son nativos de países hispanohablantes, aunque no tengan formación específica de profesor de español.

En el ámbito universitario hay varios cursos de licenciatura en lengua española. La *Universidade Federal de Pernambuco* (UFPE) fue la primera en el estado de Pernambuco que ofreció esa licenciatura en 1950. Actualmente, la UFPE ofrece un posgrado en Lingüística Aplicada a la lengua española y un curso a distancia para 150 alumnos. Además, otras instituciones de enseñanza superior también ofrecen licenciatura en Letras y preparan profesores para trabajar en la enseñanza del español (Martínez-Cachero Laseca, *op. cit.*).

Para finalizar esta breve contextualización, nos referimos a las palabras de Martínez-Cachero Laseca (2009: 194) sobre su percepción del futuro de la enseñanza del español en Pernambuco:

*É certo que a situação do espanhol em Pernambuco não é muito próspera, especialmente no ensino não universitário, contudo, a forte demanda que as atividades formativas têm; a realização de diversos Congressos; o dinamismo de sua Associação de Professores de Espanhol ou as cifras do espanhol no Vestibular, nos fazem ver com otimismo o futuro do espanhol neste Estado.*



## 7.2. La selectividad: el *vestibular*

Como mencionamos anteriormente, una de las pruebas de la creciente vitalidad del español en la enseñanza secundaria, bachillerato y en las universidades brasileñas viene dada por el número de candidatos que realizan la prueba de lengua española en el examen *vestibular*, prueba de selectividad organizada por las universidades para escoger a sus nuevos estudiantes. Cuando el alumno se inscribe en el *vestibular*, debe elegir el idioma extranjero del que será examinado, y que puede ser diferente del que cursó a lo largo de su formación escolar. En 1998 casi todas las universidades del país, federales y estatales, públicas y privadas, incluían en sus procesos de selectividad el conocimiento de español, que llegó a ser la lengua extranjera más demandada, por delante incluso del inglés en algunas universidades.

En el gráfico abajo, podemos observar la evolución del número de candidatos inscritos en el examen de lengua española en el *vestibular* de la UFPE (figura 5).

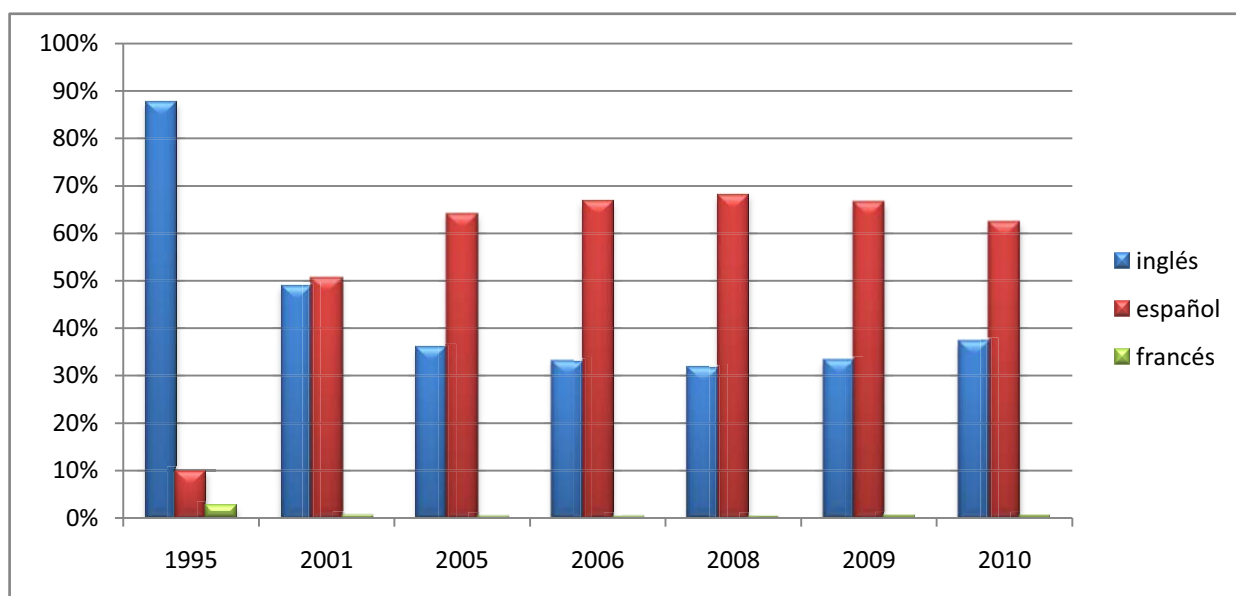


Figura 5. Número de candidatos por lengua extranjera en el *vestibular* de la UFPE.

Fuente: COVEST COPSET, 2011.

En el año de 1994 se empezó a ofrecer el español, además del inglés y del francés, como una opción más para el examen de lengua extranjera en el *vestibular* de la UFPE. En la figura 5, vemos que en 1995 el español todavía no era una opción tan procurada como el inglés. En

esa época, el inglés representaba el 87,4% del total de inscritos y poco más de 10% de los alumnos elegían el español.

En el año 2005, diez años después de la inclusión del español en el *vestibular* de la UFPE, se observa un incremento significativo en el número de alumnos que optaron por la lengua española, cuando de cada cinco candidatos inscritos en la primera etapa<sup>23</sup>, tres eligieron el español como la opción en el examen de lengua extranjera. La preferencia por hacer el examen de lengua extranjera del *vestibular* de la UFPE en lengua española se ha mantenido desde entonces (figura 5).

La justificativa que da la gran mayoría de los alumnos para su elección de la lengua española es la proximidad lingüística entre el portugués y el español. Es muy frecuente oírse la frase '*Aprender espanhol é fácil porque falo português*', como vemos reflejado en el título del artículo de Andrade Neta (2000).

Con respecto al número de candidatos al curso de Letras por el número de plazas que se ofrecen, observamos que aunque haya habido un mayor interés por el curso de Letras hasta el 2005 cuando se registraron 11 candidatos por plaza, este número ha bajado en los últimos años y en 2009 hubo una media de 5,9 candidatos por plaza (figura 6).

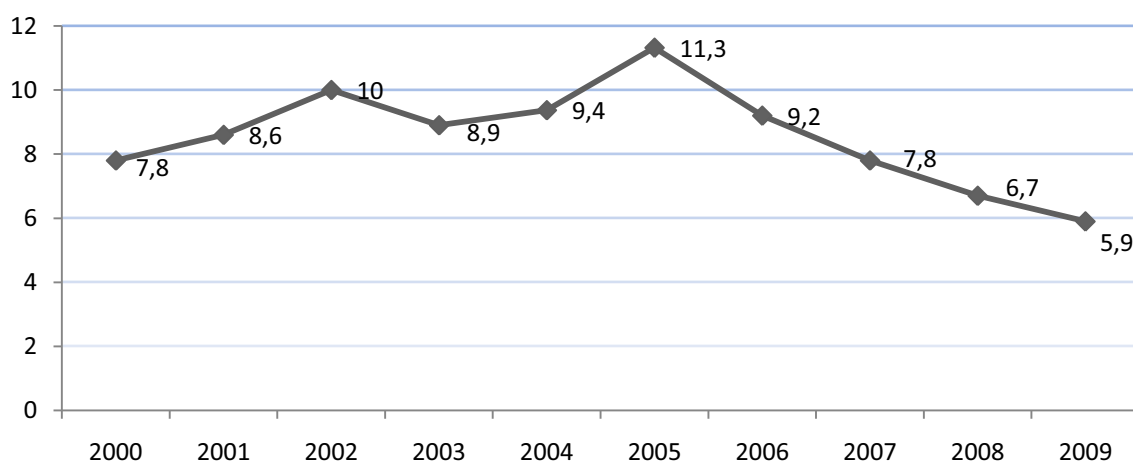


Figura 6. Número de candidatos al curso de Letras de la UFPE/nº de plazas.

Fuente: COVEST COPSET, 2011.

<sup>23</sup> El *vestibular* se divide en dos etapas. En la primera, los candidatos se examinan en todas las materias, y en la segunda, sólo en las materias específicas del curso universitario que quieren estudiar.

La disminución en el número de candidatos del curso de Letras es una tendencia contraria a las expectativas generadas con la entrada en vigor de la llamada ley del español, es decir, la ley 11.161 de 2005, que detallamos en el capítulo 7. (La enseñanza de español en Brasil). Pero como bien señala Pérez de Obanos (2010) los bajos salarios, entre otros factores, son problemas que dificultan la consolidación y el prestigio de la carrera del profesor y empuja a los jóvenes universitarios a estudiar otras profesiones.

Es importante señalar que el examen *vestibular* se está sustituyendo poco a poco por el *Exame Nacional do Ensino Médio* (ENEM). Este examen fue creado en 1998 con el objetivo de evaluar los conocimientos del estudiante al final de la educación básica. Actualmente, aproximadamente 500 universidades utilizan el resultado de este examen como criterio de selección para la entrada en la enseñanza superior, tanto complementando como sustituyendo el *vestibular* (INEP, 2011).

A continuación describimos las características del curso de Letras de la UFPE, donde recogimos los datos que forman la muestra analizada en esta investigación.

### 7.3. El curso de Letras de la *Universidade Federal de Pernambuco* (UFPE)

El curso de graduación en Letras de la UFPE se divide en *Licenciatura* y *Bacharelado*. Los alumnos que eligen la primera modalidad tienen que hacer asignaturas pedagógicas, mientras que en la segunda esto no es necesario.

Hasta el año 2009, los alumnos de Letras que optaban por estudiar una lengua extranjera seguían el currículo de la licenciatura que les otorgaba el diploma de licenciados en lengua y literatura en lengua portuguesa y en la lengua extranjera que habían estudiado (inglés, español o francés). Los alumnos que hacen parte de este estudio seguían el currículo de la licenciatura que les otorgaba el diploma de licenciados en lengua y literatura portuguesa y española (tabla 1).

Tabla 1. Programa de lengua española del curso de Letras de la UFPE.

Curso	Semestre	Asignaturas de lengua	Nivel según el MCER	Número de horas	Número de horas acumuladas
1º	1	Lengua española 3	A1	90	90
	2	Lengua española 4	A1	60	150
2º	1	Lengua española 5	B1	60	210
	2	Lengua española 6	B1	60	270
	2	Literatura I	B1	60	-
3º	1	Lengua española 7	B2	60	330
	1	Literatura II	B2	60	-
	2	Lengua española 8	B2	60	390
	2	Literatura III	B2	60	-
4º	1	Prácticas de enseñanza 1	C1	150	540
	2	Prácticas de enseñanza 2	C1	150	690

El curso se organizaba por semestres y duraba de media cuatro años, un mínimo de seis semestres y un máximo de catorce semestres. La carga horaria total del curso correspondía a 2.460 horas de las que 750 horas se dedicaban exclusivamente a la lengua española (390h), a la literatura de lengua española (180h) y a las prácticas de enseñanza de lengua española (180h). En el anexo 1 se puede consultar el currículum de los cursos.

El cuerpo docente del Departamento de Letras contaba, en el período de la recogida de datos, con cinco profesores en su plantilla, todos hablantes nativos de lengua española (tres de España, uno de Argentina) que tenían conocimiento de la variante brasileña de la lengua portuguesa (L1). Sin embargo, hacían falta más profesores para hacer frente a la demanda del propio curso de Letras además de la demanda de los cursos de Comunicación, Secretariado y Turismo de los que la lengua española también figura en el currículum.

En el 2010 el curso de Letras sufrió una reforma curricular y desde entonces los alumnos del curso de Letras tienen que optar por estudiar una sola lengua, es decir, se licencian en portugués, o inglés, o francés o español. Esto ha aumentado el número de asignaturas específicas de la lengua que se está estudiando y ha permitido que los estudiantes focalicen más en el estudio de la lengua meta.

## PARTE III: ESTUDIO EMPÍRICO

### 8. OBJETIVOS ESPECÍFICOS Y PREGUNTAS DE INVESTIGACIÓN

Este trabajo tiene como objetivo general investigar los errores léxicos en textos escritos en español por alumnos universitarios brasileños en formación como profesores de español lengua extranjera (ELE).

El objetivo general se desglosa en los siguientes objetivos específicos:

- Examinar la fluidez y la diversidad léxica en la producción escrita de alumnos universitarios brasileños de lengua española en diferentes niveles del programa de formación como profesores de ELE;
- Identificar los errores léxicos en los textos escritos;
- Conocer qué cambios se dan en la ocurrencia de los errores léxicos entre los diferentes niveles;
- Averiguar la influencia de la L1 en el origen del error en los diferentes niveles;
- Verificar si se observa una tendencia a la fosilización de los errores léxicos identificados.

Con relación a estos objetivos, formulamos las siguientes preguntas:

1. En la producción escrita de los profesores brasileños de ELE en formación ¿cambia la fluidez y la diversidad léxica a lo largo del período de formación?
2. ¿En qué categoría de palabra (léxica o funcional) se observa un mayor número de errores según el número de horas de instrucción?
3. ¿A qué clase de palabra léxica (nombre, adjetivo, verbo o adverbio) afectan los errores a medida que aumenta el número de horas de instrucción?

4. ¿Qué tipo de errores léxicos (formal, semántico o mixto) se observan en los textos escritos por profesores brasileños de ELE en formación según el número de horas de instrucción?
  
5. ¿Cuál es el origen de los errores léxicos en la producción escrita de los estudiantes brasileños de ELE a lo largo del período de formación? ¿Los errores se deben a la influencia de su L1 (error interlingüístico), el portugués variante brasileña, o son fruto del proceso de desarrollo de la interlengua (error intralingüístico)?
  
6. ¿Cambia el tipo del error léxico (formal, semántico o mixto) según el origen (interlingüístico, intralingüístico o ambiguo) en la producción escrita de los estudiantes brasileños durante el período de formación?

Para finalizar, nos ha parecido pertinente completar el estudio con una exploración de la percepción que tienen los estudiantes de la proximidad lingüística y su efecto (facilitador o no) en la adquisición del léxico, así como de la percepción que tienen de sus capacidades en las cuatro destrezas (leer, escuchar, escribir y hablar), en especial de la habilidad de expresión escrita, que presentamos en el capítulo de resultados. Sin embargo, resaltamos que este tema no hace parte de los objetivos específicos del presente estudio, ya que constituye un tema amplio y complejo que requeriría una exploración propia.

A continuación presentamos la metodología que utilizamos para desarrollar esta investigación.

## 9. METODOLOGÍA

El objetivo principal de la presente investigación es conocer el tipo y el origen del error que producen los participantes, alumnos brasileños de lengua española, en las palabras léxicas en diferentes niveles de su proceso de aprendizaje y su formación como futuros profesores de español como lengua extranjera (ELE). En el estudio también se procura explorar la tendencia a la fosilización de determinados errores. Para ello, se optó por un diseño pseudolongitudinal y la aplicación de una metodología cuantitativa para el estudio de los errores completada por una breve exploración cualitativa de aquellos errores que pudieran presentar una tendencia a la fosilización.

Se diseñó un cuestionario para obtener los datos personales de los participantes (anexo 2). Para la recogida de datos lingüísticos se propusieron diversos temas con el objetivo de estimular la producción escrita (anexo 3). También se diseñó un cuestionario en el que se formulaban preguntas sobre la percepción que tenían los participantes de su nivel de español y de la proximidad lingüística entre el portugués (L1) y el español (L2) (anexo 4).

A continuación explicamos los métodos de investigación cuantitativa que se utilizan en el campo de la Adquisición de Segundas Lenguas (ASL). Después, comentamos el diseño pseudolongitudinal que se ha escogido para llevar a cabo esta investigación, presentamos las características de los participantes del estudio, los instrumentos y procedimientos para la recogida de datos lingüísticos y su organización, así como los instrumentos para el procesamiento y análisis de los mismos.

### 9.1. Los métodos de investigación cuantitativa en ASL

Según Larsen-Freeman y Long (1991), un estudio cuantitativo típico es un experimento diseñado para comprobar hipótesis a través del uso de instrumentos objetivos y análisis estadístico adecuado. El análisis estadístico se ha mostrado un aliado importante en los estudios cuantitativos en el ámbito de ASL porque posibilita no sólo la cuantificación de la

ocurrencia de un determinado fenómeno, como la explicación de las relaciones entre variables, comprobación de la eficacia de la instrucción y generalización de los resultados<sup>24</sup>.

En un estudio cuantitativo, además de tener claro el objetivo de la investigación y la elección de la prueba estadística, es necesario comprobar si los datos obtenidos cumplen con los supuestos teóricos relacionados con cada tipo de prueba estadística (Lazaraton, 2005). Es importante destacar que hay diferentes pruebas estadísticas y que cada una de ellas se utiliza cuando se cumplen determinadas condiciones o supuestos teóricos. Se profundiza en este tema en el apartado 9.7.2. Procesamiento estadístico (SPSS), en el que se describen las pruebas estadísticas utilizadas en la presente investigación.

El uso de la investigación cuantitativa ofrece ventajas como la sistematicidad, rigurosidad, la objetividad y el control, además de que la investigación cuantitativa produce resultados fiables, replicables y que pueden ser generalizados para otros contextos (Dörnyei, 2007). Es importante destacar que esas ventajas dependen directamente del cumplimiento de los supuestos teóricos en los que se basan cada una de las pruebas estadísticas.

Como se ha dicho más arriba, la utilización de determinadas pruebas estadísticas permite la generalización de los resultados a contextos similares. Sin embargo, Lazaraton (*op. cit.*) advierte que es necesario no sólo conocer las ventajas y desventajas de la metodología que se utiliza, sino que también es importante restringir las conclusiones del estudio en el marco de las limitaciones de la metodología que se ha empleado. Esa es una crítica que se puede hacer a los métodos cuantitativos, una vez que resultados similares pueden ser el resultado de diferentes procesos subyacentes. De ser así, la utilización exclusivamente de métodos cuantitativos no permite al investigador captar la situación particular de cada uno de estos procesos.

Muchos investigadores que defienden los estudios de tipo cualitativo han hecho críticas al método cuantitativo porque lo perciben como simplista, descontextualizado e incapaz de capturar las particularidades de los individuos (Firth y Wagner, 1997). Lo ven así porque la información que el estudio de tipo cuantitativo generalmente ofrece se basa en el promedio de respuestas de todos los participantes del grupo observado, y al tener en cuenta

---

<sup>24</sup> En Loewen y Gass (2009) se puede encontrar una recopilación de los estudios que emplean la estadística en el campo de la ASL de 1948 a 2006. Los autores recogen en una tabla el objetivo principal de los estudios y el método estadístico utilizado en cada una de las investigaciones.



solamente los promedios es imposible hacer justicia a la variedad subjetiva inherente al individuo.

Pese a las críticas, los estudios cuantitativos siguen siendo muy utilizados en el ámbito de la ASL. Un ejemplo de ello se encuentra en el estudio sobre los métodos de investigación cuantitativa utilizados en el ámbito de la lingüística aplicada realizado por Lazaraton (2005). En ese estudio la autora analizó 524 investigaciones publicadas entre los años 1991 a 2001 en cuatro revistas de lingüística aplicada<sup>25</sup>. El objetivo de la investigación era conocer la metodología cuantitativa que se estaba utilizando por los investigadores de ASL. El estudio reveló que 450 (86%) de las investigaciones eran de tipo cuantitativo, 67 (13%) cualitativo y 7 (1%), mixto.

En lo que se refiere al tipo de procedimiento estadístico utilizado en la investigación en ASL, Lazaraton (2005) encontró que el procedimiento más común encontrado era el llamado descriptivo<sup>26</sup> (377; 84%), en el que se presentan las frecuencias, porcentajes, medias y desviaciones estándar. La autora encontró asimismo que el segundo procedimiento más utilizado era el test estadístico de análisis de la varianza<sup>27</sup> (199; 44%), que procura conocer el efecto entre las variables de un estudio.

Ortega y Iberri-Shea (2005) observaron que muchos estudios cuantitativos de tipo longitudinal utilizan solamente el procedimiento estadístico descriptivo y atribuyen ese hecho al número reducido de participantes en la muestra. Como se ve en el apartado 9.7.2. Procesamiento estadístico (SPSS), para utilizar la estadística inferencial<sup>28</sup> como prueba del análisis de la varianza, se deben cumplir condiciones específicas como por ejemplo, la

---

<sup>25</sup> *Language Learning, Studies in Second Language Acquisition, Modern Language Journal, and TESOL Quarterly.*

<sup>26</sup> La estadística descriptiva ofrece una visión general de los datos de la muestra. Asimismo, permite la organización y clasificación de los indicadores cuantitativos obtenidos en la medición. A partir del análisis de esos datos, se perciben relaciones y tendencias del fenómeno. La información se organiza a través de tablas de distribución de frecuencias, gráficos y las medidas de tendencia central como la mediana y la media, entre otros (Mackey y Gass, 2005).

<sup>27</sup> ANOVA, análisis de la varianza, prueba utilizada en la estadística inferencial. Se profundiza en este método de análisis estadístico en el apartado 9.7.2. Procesamiento estadístico (SPSS).

<sup>28</sup> La estadística inferencial se emplea en la interpretación y valoración cuantitativa del fenómeno que se estudia y permite al investigador la generalización de los resultados a contextos similares (Vallejo, 2008).

presencia de un número mínimo de participantes en la muestra. Ese tipo de investigación centra su atención en las características comunes a determinados individuos que se agrupan a través de la cuantificación de datos recogidos para formar una variable específica de análisis. Investigadores como Lazaraton (2005) y Dörnyei (2007) recomiendan que las variables utilizadas estén bien definidas y que los descriptores para el rango de valores que se permiten dentro de la variable sean precisos para que los números obtenidos a partir del agrupamiento de las variables tengan sentido.

Sin embargo, Loewen y Gass (2009) destacan el aumento del interés por el uso de la estadística inferencial en la investigación en ASL y por la preocupación con el rigor tanto en la utilización de los procedimientos estadísticos como en la divulgación de los resultados. Además, los autores revelan que en los últimos 10 años, el número de artículos publicados<sup>29</sup> en ASL que utilizan métodos estadísticos supera el número de artículos que no lo hacen.

En lo que se refiere a la investigación cualitativa que se desarrolla en el campo de la ASL, muchas veces ésta se refiere a estudios etnográficos en los que los investigadores se centran en observar qué ocurre en un determinado contexto (Larsen-Freeman y Long, 1991). Sin embargo, se ha notado en los últimos años que, debido a la complejidad de la investigación que se venía realizando, era necesario un abordaje más amplio que comprendiera una gran variedad de procedimientos (Chapelle y Duff, 2003). No obstante, esta variedad alberga una serie de características comunes como descripciones extensas y ricas, representación natural y holística, pocos participantes, perspectiva émica, proceso cíclico, posibilidad de adoptar una orientación ideológica y preguntas de investigación abiertas y generales (Mackey y Gass, 2005).

Muchas veces se habla de estudios cualitativos en oposición a estudios que utilizan una metodología cuantitativa. Sin embargo, durante la elaboración de la presente tesis, encontramos algunas investigaciones en ASL que utilizan ambos paradigmas de investigación. Un ejemplo de ello es el artículo de Cantos Gómez (2002), en el que el autor apoya el empleo de esas dos técnicas y demuestra con ejemplos cómo la utilización de los métodos cuantitativo y cualitativo contribuye en el análisis lingüístico y enriquecen los

---

<sup>29</sup> Artículos publicados en las revistas *Language Learning*, *Studies in Second Language Acquisition*, and *Second Language Research*.

resultados. Asimismo, hay autores como Lazaraton (2005) y Dörnyei (2007) que defienden la utilización de métodos mixtos de investigación porque la combinación de estos dos paradigmas aporta diferentes miradas sobre el mismo fenómeno, que se complementan y posibilitan una mejor comprensión de los datos. Además, así como Brown (2004) pensamos que la distinción entre los métodos cuantitativo y cualitativo debe ser vista como un *continuum* y no una dicotomía.

En definitiva, hay cada vez más investigaciones en ASL que recurren al método cuantitativo en el intento de conocer y explicar los fenómenos lingüísticos que ocurren en el proceso de aprendizaje de una L2. Además, la utilización de métodos mixtos de investigación puede ofrecer una mejor comprensión de dichos fenómenos. En lo que se refiere a la presente tesis, aunque se apoye más en los métodos cuantitativos para responder las preguntas de investigación, se ha realizado también un breve análisis lingüístico cualitativo de los errores léxicos identificados en el *corpus* con el fin de ratificar, ilustrar y enriquecer los hallazgos de las pruebas estadísticas.

## 9.2. El diseño pseudolongitudinal

La elección del método depende en gran parte de las preguntas de investigación (Mackey y Gass, 2005). En cuanto al estudio del error, la mayoría de las investigaciones empíricas realizadas siguieron una metodología de diseño transversal, es decir, estudios en los que se recopilan datos de un gran número de sujetos en un momento específico del proceso de aprendizaje con la idea de ofrecer una mirada sobre parte de su desarrollo (Gass y Selinker, 2008). Sin embargo, se suele recomendar la realización de estudios longitudinales<sup>30</sup> en los que se siguen a los mismos sujetos durante un período del proceso de aprendizaje. La idea es que a partir de los datos obtenidos en estudios longitudinales es posible seguir el desarrollo de la interlengua de los estudiantes, lo que puede revelar en qué áreas del lenguaje los errores persisten a lo largo del tiempo (Ellis, 2008). Esta información es especialmente útil en el estudio de la fosilización, como se ha visto en el capítulo 4.

---

<sup>30</sup> En los estudios longitudinales se acompaña a un individuo o un pequeño grupo de individuos a lo largo de un determinado período de tiempo. Para una aproximación a los estudios de tipo longitudinal desarrollados en ASL, se recomienda consultar el artículo de Ortega y Ibarra-Shea (2005).

En este estudio se ha optado por la realización de una investigación con un diseño pseudolongitudinal. Se ha tomado esta decisión porque las características de ese diseño se adecuan a los objetivos de la investigación, es decir, conocer los errores léxicos en los textos escritos de los profesores brasileños de ELE en diferentes niveles del programa de formación. Estos estudios consisten en investigar poblaciones distintas en diferentes niveles de competencia lingüística del proceso de aprendizaje. Son como estudios transversales, pero en lugar de recoger datos de un único grupo en momentos puntuales durante un período de su formación, participan del estudio varios grupos con competencia diferente (Gass y Selinker, 2008). Así se puede tener una visión general de lo que ocurre en el proceso de enseñanza-aprendizaje a partir de la información obtenida en cada uno de los niveles de competencia sin la necesidad de seguir al mismo grupo a medida que éste avanza en su formación.

Se puede argumentar a favor y en contra del diseño pseudolongitudinal, pero su eficacia va a depender del objetivo de la investigación. Si la investigación procura conocer qué ocurre a lo largo del programa de formación, el método pseudolongitudinal es el más indicado porque no se ocupa de los cambios a nivel del individuo, sino del grupo (Dörnyei, 2007). Asimismo, Han (2004) afirma que el método pseudolongitudinal es eficaz porque revela información sobre la construcción de la interlengua, pero matiza que si el objetivo del estudio es conocer las características individuales, una vez que cada participante del grupo presenta características propias, el método más indicado sería un estudio de tipo longitudinal.

Como ejemplo de estudios pseudolongitudinales en ASL podemos citar la tesis doctoral de Ainciburu (2007) que se dedica al estudio de la adquisición del léxico en las lenguas afines con alumnos italianos de ELE. Otro ejemplo de estudio relevante es el de Palapanidi (2010) que investiga los errores léxicos intralingüales semánticos en la expresión escrita en español por alumnos griegos utilizando el método del análisis de errores. Sin embargo, el uso del modelo pseudolongitudinal para analizar la evolución de diferentes categorías de errores léxicos en diferentes niveles lingüísticos es criticado por Buyse (2011), quien afirma que para medir la evolución lo más adecuado es la utilización del modelo longitudinal, en el que se hace un seguimiento de las mismas poblaciones.

Además de todo lo que se ha dicho arriba, creemos que el diseño pseudolongitudinal es particularmente adecuado para el presente estudio por cuatro razones. En primer lugar, el diseño pseudolongitudinal no exige un período prolongado para la recogida de datos (Dörnyei, 2007). Como los participantes de la muestra se encontraban en Brasil y la investigadora en España, el tiempo y los recursos económicos disponibles eran limitados. La segunda razón se refiere al hecho de que la utilización del método pseudolongitudinal evita que ocurra una disminución en el número de participantes a lo largo del período de recogida de datos como puede ocurrir en estudios de tipo longitudinal. Ese fenómeno se conoce como *mortalidad de los participantes (participant mortality)* (Mackey y Gass, 2005: 111). La tercera razón se debe al hecho de que ese diseño nos permite estudiar un grupo más representativo de individuos, es decir, más numeroso, algo que no es viable en estudios de tipo longitudinal. Además, los estudios pseudolongitudinales permiten, muchas veces, la generalización de los resultados a poblaciones con características similares (Alba Quiñones, 2009). Finalmente, el diseño pseudolongitudinal disminuye la predisposición del participante a modificar su respuesta natural por la interferencia que el contacto prolongado con los observadores o instrumentos de un estudio puede producir, como en los estudios de tipo longitudinal (Shadish *et al.*, 2001).

### 9.3. Participantes

En nuestro estudio colaboraron en total 142 alumnos de las asignaturas de lengua española del Departamento de Letras de la *Universidade Federal de Pernambuco (UFPE)*, Brasil. En nuestra muestra hay un predominio de mujeres, como puede verse en la tabla 2 y la mayoría de los participantes se encuentra en la franja de edad entre 20 y 30 años (figura 7).

Tabla 2. Género de los participantes.

	N	%
Mujeres	114	80,3
Hombres	28	19,7
Total	142	100

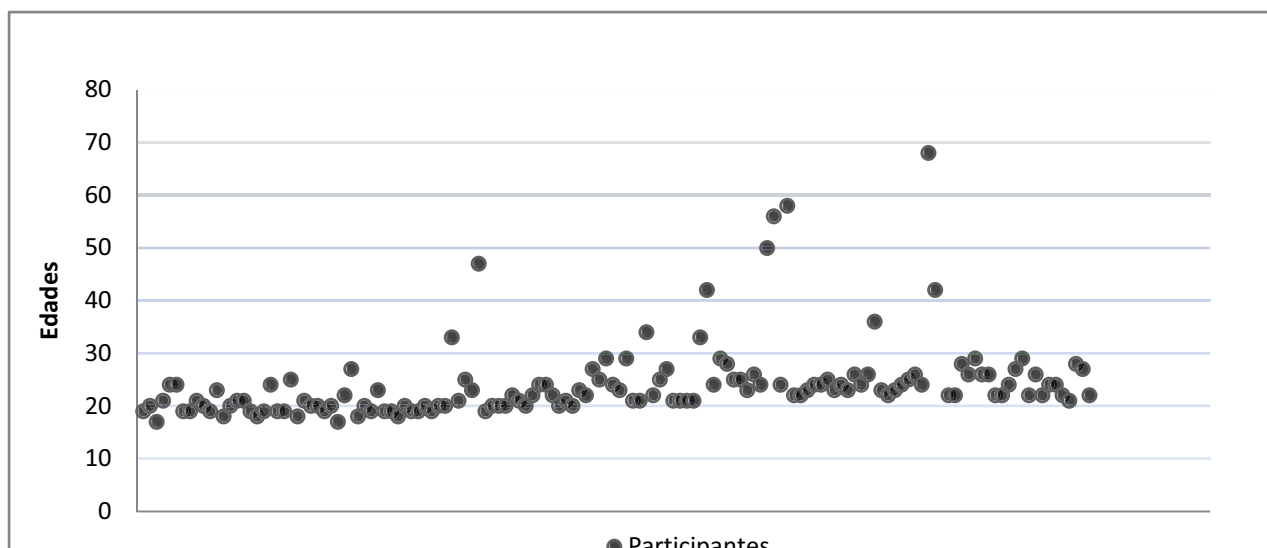


Figura 7. Edad de los participantes de la muestra.

Tal como revela la información obtenida a través del cuestionario de datos personales (anexo 2) la L1 de los participantes es el portugués, variedad brasileña, y la mayor parte de ellos (N=134; 94%) nació en Pernambuco, un estado de Brasil que no hace frontera ni está cerca de ningún país hispanohablante, por lo que la influencia del uso del portugués<sup>31</sup> es mínima. Los demás participantes (N=8; 6%) provienen de los estados de Ceará, São Paulo y Mato Grosso do Sul. La situación geográfica de Ceará es similar a la de Pernambuco en lo que se refiere a la inexistencia de fronteras con países hispanohablantes. São Paulo es un estado de Brasil que es muy cosmopolita y recibe a más visitantes de otros países que los demás. Aunque el estado de Mato Grosso do Sul hace frontera con Bolivia y Paraguay, lo que sugiere que pueda haber un mayor contacto de sus habitantes con individuos hispanohablantes (Martínez-Cachero Laseca, 2008), con excepción de casos específicos, no se puede generalizar la presencia del español en la región.

El 86% de los participantes tuvo su primer contacto con el idioma en un contexto educativo, en el instituto o en cursos preparatorios para el *vestibular* (prueba de selectividad organizada por las universidades para escoger a sus nuevos estudiantes) a partir de los 15 años de edad. La mayor parte de los estudiantes (92%) realizó los exámenes de selectividad a partir de 1994, año en el que el español pasó a figurar como una opción más en el examen de lengua extranjera. De entre esos alumnos, el 91,7% prefirió hacer ese examen en lengua

<sup>31</sup> Véase apartado Lenguas afines: el portugués y el español.

española y el 8,3% restante lo hizo en inglés porque consideraban tener un mayor dominio de ese idioma.

De los participantes que colaboraron en el estudio, apenas el 7% había tenido la oportunidad de viajar a un país hispanohablante, siendo el período máximo de la estancia inferior a una semana. Dado el corto período de estancia en el extranjero, consideramos que la experiencia de contacto de esos participantes con la lengua meta en un entorno hispanohablante no impactó de manera significativa en su habilidad de expresión escrita, por lo que no excluimos las producciones de esos alumnos de la muestra, con excepción de un único caso en el que el participante había estudiado durante un año en Salamanca, España.

En lo que se refiere al conocimiento de otras lenguas, el 57% de los participantes afirmó no conocer otra lengua extranjera aparte del español. Para el 43% restante, el inglés, seguido del francés, eran los idiomas extranjeros que más conocían los alumnos. Eso se debe al hecho de que el inglés es el idioma extranjero obligatorio en la enseñanza secundaria y que el francés sigue siendo un idioma de prestigio en Brasil (MEC, 2009).

Un poco más de la mitad de los participantes (51%) afirmó practicar el español fuera del aula. Las actividades más habituales cuando no están en clase son: charlas con los compañeros de curso (55,6%), conversaciones a través de *chat*/internet (30,6%) con otros hablantes no nativos de español y conversaciones con hablantes nativos (18,1%). Asimismo, los participantes revelan que procuran acercarse a hablantes nativos de lengua española siempre que posible por el interés que tienen en aprender la lengua y la cultura, aunque la presencia de hispanohablantes no es muy común en esta región de Brasil<sup>32</sup>, aunque se observe un crecimiento entre los años 2008 y 2009.

<sup>32</sup> Turismo receptivo en Pernambuco en 2008 y 2009.

Continentes y países de residencia permanente	N de turistas procedentes de países hispanohablantes	
	2008	2009
América Central y Caribe	67	127
América del Norte: México	53	135
América del Sur	6.125	9.544
Europa: España	2.929	4.550
Total	9.174	14.356

Fuente: Secretaria Nacional de Políticas de Turismo – Brasil (2010).

#### 9.4. Instrumento de recogida de datos lingüísticos

Para la recogida de datos lingüísticos que constituye nuestro *corpus* de análisis, se diseñaron pruebas escritas (anexo 3) en las que se solicitaba a cada alumno escribir tres textos sobre tres temas distintos previamente seleccionados (Asselin y McLaghlin, 1992)<sup>33</sup>.

Se tomó la decisión de que los alumnos deberían escribir tres textos con la finalidad de lograr un *corpus* más amplio, ya que habíamos observado durante la realización del estudio exploratorio que se realizó para la obtención del Diploma de Estudios Avanzados (DEA) que los estudiantes solían redactar textos muy breves, pese a la instrucción que indicaba el número mínimo de palabras solicitado para cada texto. Una gran variación en el tamaño de los textos producidos por los participantes también se ha observado en otras investigaciones (ej.: Lehmann, 2007).

Los criterios que seguimos para escoger los temas fueron que todos los participantes tuvieran algo que transmitir a partir de sus propias experiencias personales, así como de sus vivencias como estudiantes de español. Además, pretendimos que los temas fueran motivadores para los alumnos y que estimularan la producción de textos con significado relevante tanto para los alumnos como para la investigadora. A continuación explicamos cada uno de los temas propuestos.

- Redacción 1

El tema propuesto para la redacción 1 se titula 'El español y tú'. Al elegir este tema, que es común a todos los participantes, quisimos no sólo conseguir un *corpus* para el análisis de errores en contexto, sino también conocer las experiencias previas de los alumnos, su trayectoria en el contacto con la lengua española y su visión personal respecto a esta vivencia.

---

<sup>33</sup> Asselin y McLaghlin (1992) utilizaron un instrumento de recogida de datos similar en su investigación. Los autores analizaron un corpus formado por 376 producciones escritas sobre temas variados y diferentes tipos de texto. El objetivo de su estudio era analizar el nivel de competencia escrita de los estudiantes de primer curso universitario de francés L2, a través de un recuento y categorización de los errores lingüísticos encontrados, y proponer criterios de evaluación con el fin de mejorar dicha competencia.



- Redacción 2

Para la redacción 2 se formularon tres temas distintos. Eso se debió al hecho de que al final de cada recogida de datos se observaba si los textos producidos cumplían con las consignas de tema y número de palabras (en el apartado 9.5. 'Procedimiento de recogida de datos', se detalla información sobre los períodos y las etapas en las que se realizaron las recogidas de datos). A continuación explicamos cada uno de los temas propuestos.

- Redacción 2a: Las fiestas de San Juan

La idea inicial para esa redacción era que los estudiantes escribieran un texto descriptivo. Al elegir el tema 'Las fiestas de San Juan' pretendíamos que los estudiantes describieran las celebraciones, la comida, la música y los bailes típicos del Nordeste de Brasil que tienen lugar durante el mes de junio. Sin embargo, al analizar las producciones del estudio exploratorio, verificamos que los textos producidos sobre ese tema contenían muchas palabras en portugués, porque hacían referencia directa a los nombres de los bailes y de las comidas típicas, que no tienen, en muchos casos, un equivalente correspondiente en español (ej.: *farró*, baile típico; *canjica*, plato típico). Algunos alumnos hicieron un intento de explicar esos elementos, pero otros simplemente usaban las palabras en portugués sin describirlas ni explicarlas. Como verificamos que ese tema no estimulaba la producción de L2 y debido a que el objetivo del presente trabajo es el estudio de los errores léxicos, decidimos proponer otro tema en la siguiente recogida de datos.

- Redacción 2b: La Ley 11.161/2005

El tema siguiente se basaba en la ley 11.161/2005 (cuyo propósito se describe en el capítulo 7. 'La enseñanza de español en Brasil') y procuraba que los alumnos expresaran su opinión respecto a esa ley y cómo veían su futuro profesional a partir del cambio que supondría la entrada en vigor de dicha ley: la oferta obligatoria de la materia lengua española en los tres últimos años de la enseñanza media. Una breve descripción de la ley fue introducida en el cuerpo de la pregunta para garantizar que todos los participantes fueran capaces de escribir sobre el tema. Sin embargo, muchos

participantes se limitaron a decir “sí, me gusta la ley”, o “sí, me parece una iniciativa positiva”, pero no explicaban de qué manera eso les afectaría y qué cambios esta nueva ley ofrecía en el panorama de la enseñanza del español para los profesores recién licenciados. Por ese motivo, se decidió proponer un nuevo tema.

- Redacción 2c: Los problemas de tu barrio

Al proponer el tema ‘Los problemas de tu barrio’ quisimos que los alumnos hablaran de sí mismos y describieran su entorno y situaciones comunes de su vida diaria en una carta al presidente del país. En la carta deberían felicitar al presidente por su buena gestión, si así les parecía, y también pedirle soluciones a los problemas específicos de su entorno. Ese tema se mostró ser más interesante que el anterior para los participantes y les ayudó a cumplir con la consigna de escribir entre 150 y 200 palabras.

- Redacción 3

El tema propuesto para la tercera producción escrita fue ‘Un lugar especial de tu infancia o juventud’. Ese tema procuraba que los participantes pudieran escribir un relato en pasado sobre algo cercano a ellos y además sobre una experiencia ya vivida. Se esperaba que los estudiantes describieran un poco el contexto y usaran las formas verbales de pasado. Este tema también obtuvo una buena respuesta de los participantes, lo que refleja que cuando el alumno se siente familiarizado con el tópico, esto le aporta seguridad y consecuentemente estimula la escritura.

## 9.5. Procedimiento de recogida de datos

La recogida de los datos que forman el *corpus* de esta investigación se llevó a cabo durante dos años en los cursos 2005/2, 2006/2, 2007/1 y 2007/2 en la ciudad de Recife, Brasil. En la *Universidade Federal de Pernambuco* (UFPE) los cursos son semestrales. Así, cuando mencionamos el curso 2005/2, nos estamos refiriendo al curso que se realizó en el segundo semestre del año 2005. Para tener acceso a los alumnos, se contactó previamente a los profesores de las asignaturas de lengua española en el Departamento de Letras de la UFPE en Brasil para informales de las características de la investigación y solicitar su colaboración.

Los profesores de seis asignaturas se mostraron receptivos y permitieron que la recogida de datos se realizara en sus horarios de clase. Como se ha detallado previamente en el apartado 7.3. (El curso de Letras de la *Universidade Federal de Pernambuco* (UFPE)), la formación específica en lengua española consta de un total de seis asignaturas de lengua, tres de literatura y dos prácticas de enseñanza.

En la tabla 3 se puede ver el número de textos obtenidos por tema propuesto y asignatura. No fue posible recoger datos de los estudiantes de Lengua Española 5 y Prácticas de Enseñanza I porque no se pudo contactar con los docentes de las respectivas asignaturas.

Tabla 3. Número de textos escritos obtenidos por asignatura en cada recogida de datos.

Asignaturas	Temas propuestos para estimular la producción de textos	N de textos obtenidos								N total de textos
		LE 3	LE 4	LE 5	LE 6	LE 7	LE 8	P 1	P 2	
Curso 2005/2	Redacción 1: El español y tú	-	-	-	9	-	-	-	12	21
	Redacción 2a: La fiesta de San Juan	-	-	-	7	-	-	-	6	13
	Redacción 3: Un lugar especial de tu infancia o juventud	-	-	-	9	-	-	-	9	18
Curso 2006/2	Redacción 1: El español y tú	16	13	-	9	7	8	-	7	60
	Redacción 2b: Las ventajas y desventajas de la ley 11.161	16	13	-	9	7	8	-	7	60
	Redacción 3: Un lugar especial de tu infancia o juventud	16	12	-	8	7	8	-	7	58
Cursos 2007/1 y 2007/2	Redacción 1: El español y tú	-	15	-	4	-	-	-	36	55
	Redacción 2c: Los problemas de tu barrio	-	14	-	4	-	-	-	26	44
	Redacción 3: Un lugar especial de tu infancia o juventud	-	8	-	4	-	-	-	24	36
<b>Total de textos por asignatura</b>		<b>48</b>	<b>75</b>	<b>-</b>	<b>63</b>	<b>21</b>	<b>24</b>	<b>-</b>	<b>134</b>	<b>365</b>

LE = Lengua Española, P= Prácticas de Enseñanza

El procedimiento seguido fue el mismo en todos momentos. Los profesores de cada asignatura informaron a los alumnos que se iba a realizar un estudio sobre el aprendizaje de español y les solicitaron su participación, que era totalmente voluntaria y anónima. Los profesores también orientaron a los participantes sobre la realización de la tarea, es decir, tras rellenar el cuestionario de datos personales, los alumnos tenían que desarrollar tres

textos y para ello disponían de dos horas. No estaba permitido consultar los diccionarios, a los compañeros, ni al profesor. Al finalizar las producciones escritas, los participantes completaban un cuestionario con breves preguntas sobre la percepción de su competencia en español, de la semejanza entre el portugués y el español, y de la dificultad para realizar la tarea de escribir en español.

## 9.6. Organización del *corpus* lingüístico

Con el fin de responder las preguntas de investigación planteadas en este trabajo y facilitar el análisis del *corpus* lingüístico obtenido, decidimos agrupar las 365 producciones escritas por los 142 participantes en tres bloques (véase tabla 4 más adelante). Para ello, seguimos tres criterios. El primero se basó en el número de horas de lengua española que cada participante había estudiado en el curso de Letras de la UFPE hasta el momento de la recogida (véase tabla 1).

El segundo se basó en el contenido lingüístico y en los objetivos específicos de las asignaturas (en el anexo 1 se puede ver el sumario de las asignaturas del curso de Letras)<sup>34</sup> en relación con los descriptores del Marco Común Europeo de Referencia (MCER, 2002) para las lenguas y del Plan Curricular del Instituto Cervantes (Instituto Cervantes, 2006). La organización del *corpus* siguiendo los descriptores del MCER, cuando ése se obtiene de participantes con diferente nivel de competencia, es una estrategia ampliamente utilizada por los investigadores (ej.: Naranjo Hernández, 2009; Palapanidi, 2010). Como se ha dicho anteriormente, el MCER<sup>35</sup> es una herramienta muy útil a la hora de comparar las habilidades lingüísticas de los alumnos con variado nivel de competencia, o procedente de diferentes instituciones.

---

<sup>34</sup> Paralelamente a los objetivos específicos señalados en el anexo 1, en cada asignatura de lengua española se trabajan también los contenidos lingüísticos correspondientes a los seis niveles identificados en el MCER (A1, A2, B1, B2, C1, C2) que presenta el Plan Curricular del Instituto Cervantes a partir de escalas de descriptores que se refieren a lo que el alumno es capaz de hacer según su nivel de competencia en lengua española.

<sup>35</sup> Aunque fue bien recibido cuando surgió, el MCER recibió críticas por la complejidad de su texto, la fiabilidad e incluso se cuestionó la necesidad de su existencia (Figueras, 2008). Sin embargo, el hecho de que permite establecer una correspondencia entre los cursos y certificaciones de diferentes instituciones en el mundo, representa una gran ventaja para todos los que se dedican a la enseñanza de L2 y por todo ello, el MCER se convirtió en un “gran dinamizador de la evaluación de lenguas” (Puig, 2008: 78).

El tercer criterio que justifica la distribución de los datos recogidos en tres bloques se basó en los resultados del análisis estadístico de la varianza<sup>36</sup> del número de errores léxicos identificados en cada uno de los textos que componen la muestra. Los resultados de las pruebas estadísticas de Kruskal-Wallis revelaron que los textos agrupados en cada bloque no presentaban diferencia significativa, lo que corroboró nuestra decisión (Bloque 1: valor-p = 0,596; Bloque 2: valor-p = 0,059; Bloque 3: valor-p = 0,753).

En la tabla 4 vemos el número de participantes por asignatura, el número de textos producidos y también como están organizados los tres bloques de análisis a partir del *corpus* lingüístico. El bloque 1 está constituido por 123 textos escritos por los 45 participantes de las asignaturas de lengua española LE3 y LE4<sup>37</sup>, correspondientes a un nivel inicial (A1 y A2 según el MCER), y que habían cursado entre 90 y 150 horas de lengua española en el curso de Letras. El bloque 2 está formado por los 108 textos de los 38 participantes de las asignaturas de lengua española LE6, LE7 y LE8, correspondientes a un nivel intermedio (B1 y B2 según el MCER), y que habían cursado entre 270 y 390 horas de lengua española en el curso de Letras. Finalmente, el bloque 3 está formado por los 134 textos de los 59 participantes de la asignatura de prácticas de enseñanza P2, correspondiente a un nivel avanzado (C1 según el MCER), y que habían cursado hasta 690 horas de lengua española en el curso de Letras.

---

<sup>36</sup> En el apartado 9.7.2. Procesamiento estadístico (SPSS) se explican las pruebas estadísticas utilizadas en el estudio.

<sup>37</sup> En el apartado 7.3. El curso de Letras de la Universidade Federal de Pernambuco (UFPE) se explican las materias de lengua española que forman el curso de Letras de la UFPE y en el anexo 1 se encuentra una copia del currículum del referido curso.

Tabla 4. Recogida de datos y formación de bloques de análisis.

Asignatura	Participantes	Redacción 1	Redacción 2 (2a, 2b, 2c)	Redacción 3	Número de textos	N total de textos
Lengua española 3 (A1)	16	16	16	16	48	N=123 Bloque 1 Nivel inicial
Lengua española 4 (A2)	29	28	27	20	75	
Lengua española 5 (B1)	-	-	-	-	-	
Lengua española 6 (B1)	23	22	22	19	63	N=108 Bloque 2 Nivel intermedio
Lengua española 7 (B1)	7	7	7	7	21	
Lengua española 8 (B2)	8	8	8	8	24	
Prácticas de enseñanza 1 (C1)	-	-	-	-	-	
Prácticas de enseñanza 2 (C1)	59	55	42	37	134	N=134 Bloque 3 Nivel avanzado
<b>Total</b>	<b>142</b>	<b>136</b>	<b>122</b>	<b>107</b>	<b>365</b>	<b>N Total textos=365</b>

### 9.7. Instrumentos para el procesamiento y análisis de los datos lingüísticos

El presente estudio se ha basado principalmente en la metodología cuantitativa para contestar las preguntas de nuestra investigación. Con ello pretendemos ofrecer una visión general de los errores en la producción escrita de los profesores brasileños de ELE en formación. Sin embargo, también se ha llevado a cabo un análisis cualitativo de los errores identificados con el objetivo de profundizar en los resultados obtenidos a partir de las pruebas estadísticas y poder matizar estos hallazgos.

Tras la recogida de datos, se asignó un código a cada texto para salvaguardar la identidad de los participantes y mantener la confidencialidad de los datos. Ese código también incluye información que identifica la asignatura que cursaba el participante, el tema de la redacción y el momento de recogida de los datos. Después se transcribió cada manuscrito a archivos de texto (formato .txt).

Con el fin de facilitar el procesamiento y análisis de los datos, se utilizaron herramientas informáticas que describimos a continuación.

### 9.7.1. El *Wordsmith Tools* y sus aplicaciones

Para codificar y analizar los datos lingüísticos utilizamos el analizador léxico *Wordsmith Tools 5* (Scott, 2010). El *WordSmith Tools 5* es un programa de análisis de *corpus* textuales. Ese programa se utiliza para facilitar el tratamiento de los datos lingüísticos porque permite explorar grandes conjuntos de textos (Gómez Palou, 2006) y analizarlos desde una perspectiva de frecuencias y concordancias de unidades léxicas (Venegas, 2008) teniendo en cuenta el contexto lingüístico.

Entre las aplicaciones que hacen parte del analizador léxico *Wordsmith Tools 5*, utilizamos el *Concord* y el *Wordlist*.

- El *Concord*

El *Concord* es una aplicación del analizador léxico *Wordsmith Tools 5* (Scott, 2010) que sirve para codificar textos y para generar listados de concordancias. Al utilizar esa herramienta es posible identificar, en un determinado *corpus*, cómo una palabra se relaciona con las demás. En este trabajo se utilizó la aplicación *Concord* para etiquetar los errores en las producciones escritas y obtener luego un listado de los errores. Se utilizó el *Concord* también para identificar los errores léxicos y examinarlos en contexto en el análisis lingüístico de tipo cualitativo.

Primeramente establecimos los códigos en base a los errores léxicos, una vez que era ése el principal objeto de nuestra investigación, tal como orienta Castillejos López (2009). En segundo lugar, etiquetamos los errores en los textos de acuerdo con los códigos previamente establecidos, siguiendo las fases de codificación que mostramos en la figura 8 abajo.

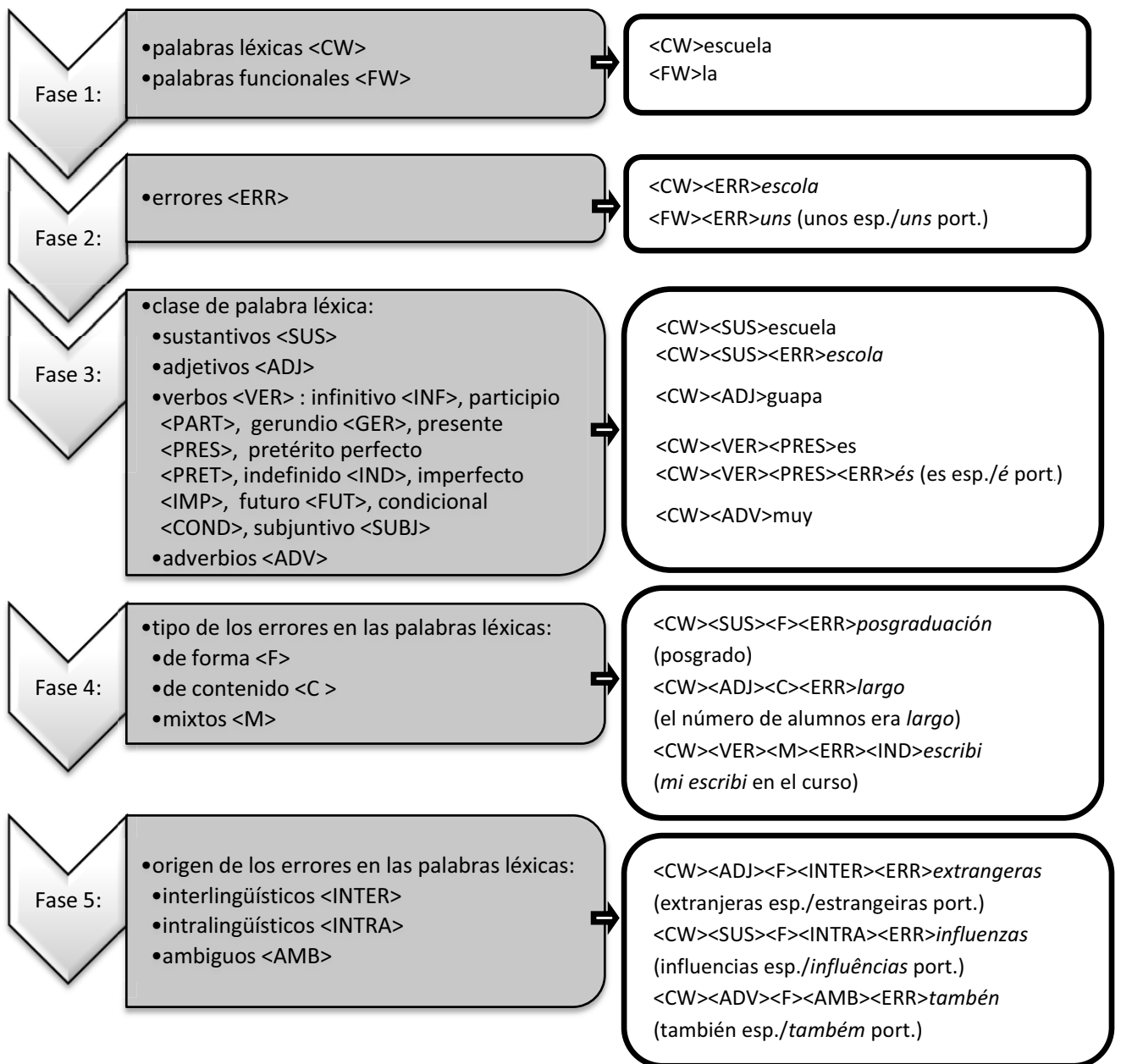


Figura 8. Fases de codificación de los datos con ejemplos.

Los códigos se insertaron entre los signos menor que y mayor que (< >), en letra mayúscula, siempre a la izquierda de la palabra a la que se refería y sin espacio entre el código y la palabra. En la primera fase se etiquetaron 43.476 palabras que forman el *corpus* lingüístico en palabras léxicas o palabras contenido (<CW> *content word*) y funcionales (<FW> *function word*). Como se ha explicado en el apartado 3.2. El error léxico, las palabras funcionales son aquellas palabras correspondientes a categorías gramaticales cerradas como preposiciones, conjunciones, pronombres, etcétera, y las palabras contenido son aquellas correspondientes a categorías gramaticales abiertas: sustantivo, adjetivo, verbo y adverbio.



En un segundo momento se marcaron todas las palabras que tenían algún tipo de error (<ERR>). En la tercera fase se identificaron y se señalaron las categorías de las palabras léxicas (<SUS> sustantivo, <ADJ> adjetivo, <VER> verbo y <ADV> adverbio), una clasificación que fue utilizada por varios investigadores como James, 1998; Celaya y Torras, 2001; Cenoz, 2001; Navés *et al.*, 2005; Agustín Llach, 2011a. En la cuarta fase, se etiquetaron los errores en las palabras léxicas según el tipo, es decir, si se trataba de un error formal (<F>), semántico (<C>) o mixto (<M>) (ej.: Fernández, 1997; James, 1998). En la quinta y última fase, los errores se etiquetaron siguiendo el criterio etiológico, es decir, si el error tenía un origen interlingüístico (<INTER>), intralingüístico (<INTRA>) o ambiguo (<AMB>) (ej.: Vázquez, 1991; Santos Gargallo, 1993; Fernández, 1997; Sánchez Iglesias, 2003; Palapanidi, 2010). En el anexo 5 mostramos un ejemplo de una redacción codificada.

Una vez realizado el etiquetado de todos los errores identificados en los textos, era posible acceder a un criterio de búsqueda determinado a través de la aplicación *Concord* (Castillejos López, 2009). Así, el investigador puede acceder con facilidad a la información que le interesa analizar, en nuestro caso, los errores léxicos. Un ejemplo de ello se puede ver en la figura 9 que ilustra la pantalla de la aplicación *Concord* en la que aparece en azul un listado de las palabras léxicas (<CW>) que presentan algún error<sup>38</sup>. La columna ‘*Concordance*’ (concordancia) incluye las palabras léxicas precedidas por la codificación correspondiente a cada una de las fases de etiquetado de los errores: la clase de palabra (<SUS>, <ADJ>, <VER>, <ADV>), el tipo (<F>, <C>, <M>) y origen del error (<INTER>, <INTRA>, <AMB>), así como de la indicación del error (<ERR>). La marca <PORT> que se observa a la izquierda de algunas palabras se ha utilizado para indicar que este ítem se escribe exactamente de la misma forma en la L1 (portugués) y que no sufrió ningún tipo de adaptación a la L2 (español). A la derecha de la pantalla se ve la columna ‘*File*’ (archivo), que indica el código atribuido a cada redacción/archivo y que contiene información sobre el tema de la redacción, la asignatura que cursaba el participante y la fase de recogida de los datos. En la línea 1 (columna N), por ejemplo, tenemos el código P2\_05\_2.3.txt que corresponde a la

---

<sup>38</sup> Dada la subjetividad asociada a la clasificación del origen del error, contamos con la ayuda de dos profesoras hablantes de español (L1), una de la variedad peninsular y otra de la variedad hispanoamericana. Además, consultamos el *Diccionario de la Lengua Española* (RAE, 2001), el *Gran Diccionario de Uso del Español Actual* (Sánchez Pérez, 2001) el diccionario *SEÑAS* (Universidad de Alcalá de Henares, 2001), el diccionario *REDES* (Bosque, 2004).

asignatura de Prácticas de Enseñanza II (P2), participante número seis (05), redacción número dos en la tercera recogida de datos (2.3).

N	Concordance	File
1	<CW><SUS>lucro <FW><ART>una <CW><SUS>parte <CW><VER><F><INTER><ERR><PORT>iria	P2_05_2.3.txt
2	<FW><PREP>de <CW><SUS>obra <FW><PREP>a <CW><SUS><F><INTER><ERR><PORT>custo	P2_05_2.3.txt
3	<FW><PREP>en 4º <CW><SUS>período , <FW><CONJ>y <CW><VER><F><INTER><ERR><GER><PORT>estudando	E6_08_1.0.txt
4	<FW><DEM>este <CW><SUS>idioma <CW><ADV>nunca <CW><VER><F><INTRA><ERR><FUT>tendra <FW><ART>un	E8_07_1.1.txt
5	<FW><INDEF>tantos <CW><SUS>pordioseros , <CW><VER><F><INTRA><ERR><GER>pediendo	P2_11_2.2.txt
6	<FW><PRO>ellos , <FW><ART>los <CW><SUS>detentos <CW><VER><F><INTER><ERR><PORT>poderian	P2_05_2.3.txt
7	<CW><SUS>posibilidades <FW><PREP>de <CW><ADV>más <CW><SUS><F><INTER><ERR><PORT>empregos .	P2_05_2.3.txt
8	<CW><ADJ>experiente <FW><CONJ><ERR>e <CW><VER><F><INTRA><ERR><IND>obtube	E6_08_1.0.txt
9	, <FW><ART>el <CW><SUS>estado , <FW><ART>las <CW><SUS><F><INTER><ERR><PORT>familias	P2_05_2.3.txt
10	<FW><PREP>del <CW><ADJ>mismo <CW><SUS><F><INTER><ERR><PORT>jeito .	P2_06_2.0.txt
11	<CW><ADJ>calurosa <CW><SUS>relación ; <CW><VER><F><INTRA><ERR><COND><PORT>seria	P2_04_1.0.txt
12	<FW><DEM>este <CW><SUS>curso , <FW><AUX>he <CW><VER><F><INTRA><ERR><PART>eligido	E6_08_1.0.txt
13	<FW><CONJ>que <CW><VER><PRES>estoy <CW><VER><F><INTER><ERR><GER><PORT>aprendendo ,	E4_09_1.1.txt
14	. <CW><VER><F><INTRA><ERR><IMPER>Aprovechala , <CW><VER><M><AMB><ERR><PRES>vas <FW><PREP>a	E6_01_3.0.txt
15	<FW><ART>una <CW><SUS>bomba <FW><PRO>se <CW><VER><F><INTRA><ERR><SUBJ>caya	P2_06_2.0.txt
16	<CW><ADJ>parada <CW><ADV>después <FW><PREP>de <CW><VER><M><INTER><ERR><INF>graduarme .	P2_06_1.3.txt
17	<CW><SUS>mes <FW><PREP>de <CW><SUS>junio . <CW><VER><F><INTRA><ERR><IMPER>Aprovechala ,	E6_01_3.0.txt
18	. <FW><PRO>lo <FW><CONJ>que <FW><PRO>me <CW><VER><C><INTER><ERR><PRES>contenta	E6_01_1.0.txt
19	<CW><ADV>más <CW><ADJ>valoradas . <FW><PRO>ellos <CW><VER><F><INTER><ERR><PRES>trabajam	P2_11_2.2.txt
20	<CW><SUS><C><INTER><ERR>solicitaciones <CW><VER><F><INTRA><ERR><SUBJ>seyan	E4_02_2.3.txt
21	<CW><ADJ><F><INTER><ERR><PORT>siguiente , 2003, <CW><VER><F><INTRA><ERR><IND>comenzé	E6_08_1.0.txt
22	<CW><SUS>salud <FW><CONJ>y <FW><ART><ERR>la <CW><SUS><F><INTER><ERR><PORT>fome .	P2_10_2.2.txt
23	<CW><SUS>partidos <FW><CONJ><ERR>e <CW><ADV>no <CW><VER><F><INTER><ERR><PORT>representam	E6_08_2.0.txt
24	. <CW><VER><PRES>es <FW><ART>un <CW><SUS><F><INTER><ERR><PORT>espetáculo	P2_15_3.0.txt
25	. <CW><ADV>así <FW><ART>los <CW><SUS>presos <CW><VER><F><INTRA><ERR><FUT>trabajaran ,	P2_05_2.3.txt
26	<CW><ADV>donde <FW><INDEF>todos <FW><PRO>se <CW><VER><F><INTER><ERR><PRES>diverten ,	E6_01_3.0.txt
27	<CW><SUS>curso . <FW><PREP><ERR>en <CW><SUS>año <CW><ADJ><F><INTER><ERR><PORT>siguiente , 2003,	E6_08_1.0.txt
28	<CW><ADV>allí <CW><VER><PRES>pienso <FW><PREP>en <CW><VER><C><INTRA><ERR><INF>confirmar	P2_04_1.0.txt
29	, <FW><ART>los <CW><SUS>trabajos <CW><ADJ><F><INTER><ERR><PORT>forçados .	P2_05_2.3.txt
30	<FW><PREP>de <CW><SUS>español <FW><PREP>para <CW><VER><C><INTRA><ERR><INF>sacar <FW><ART>el	E6_01_1.0.txt
31	<FW><CONJ>e <CW><SUS>periódicos <FW><CONJ>e <CW><VER><M><INTER><ERR><PRES>son	E6_08_2.0.txt

Figura 9. Ejemplo de pantalla de la aplicación *Concord*.

La aplicación *Concord* también permite acceder al texto completo al hacer doble clic en cualquiera de las líneas del listado. De esa forma, pudimos editar el etiquetado de los textos y ampliar la información sobre el contexto en el que ocurría un determinado error.

- El *Wordlist*

El *Wordlist* es otra aplicación del *Wordsmith Tools 5* (Scott, 2010) y sirve para visualizar listados de las palabras utilizadas en los textos y su frecuencia en el *corpus*. Se pueden obtener listados alfabéticos o por orden de frecuencia.

Una vez realizado el etiquetado de los datos con la ayuda de la aplicación *Concord* como se ha descrito en el apartado anterior, el siguiente paso era realizar la cuantificación de los datos codificados. Para ello utilizamos el *Wordlist*. Esta aplicación genera tablas con los

datos estadísticos de la muestra y provee información como el número total de *tokens* en cada texto, por ejemplo. Así, a través de esta aplicación pudimos obtener listas de frecuencia en las cuales figuraban el número de *types*, *tokens* y las ratios de aquellos elementos que nos interesaban buscar en cada uno de los textos, es decir, los errores léxicos. En la

figura 10 se observa un ejemplo de la pantalla de la aplicación *Wordlist* referente al número de errores interlingüísticos. En la pantalla se observa el código de cada redacción en la segunda línea (*text file*). El código E3\_01\_1.1 txt indica que el texto 1 (N1) de la asignatura Lengua Española 3 (E3), participante 1 (01), redacción 1 en la primera recogida de datos (1.1) presenta tres errores interlingüísticos (*tokens used for wordlist*), mientras que el texto 2 (N2), código E3\_02\_1.1 txt, y el texto 3 (N3), código E3\_03\_1.1 txt, también producidos por alumnos de la asignatura Lengua Española 3, redacción 1, en la primera recogida de datos, presentan respectivamente nueve y dos errores léxicos interlingüísticos (*tokens used for word list*).

N	Overall	1	2	3
text file		E3_01_1.1.txt	E3_02_1.1.txt	E3_03_1.1.txt
file size	320.229	2.009	1.735	1.864
tokens (running words) in text	18.498	112	94	109
tokens used for word list	0	3	9	2
sum of entries	0	0	0	0
types (distinct words)	373	3	6	2
type/token ratio (TTR)		100,00	66,67	100,00
standardised TTR				
standardised TTR std.dev.				
standardised TTR basis	1.000	1.000	1.000	1.000
mean word length (in characters)	9,40	8,67	8,22	8,50
word length std.dev.	2,43	3,51	3,15	0,71
sentences	167	1	2	1
mean (in words)	3,41	3,00	4,50	2,00
std.dev.	2,31		2,12	
paragraphs	130	1	1	1
mean (in words)	4,38	3,00	9,00	2,00
std.dev.	2,68			
headings	130	1	1	1
mean (in words)	4,38	3,00	9,00	2,00
std.dev.	2,68			
sections	130	1	1	1
mean (in words)	4,38	3,00	9,00	2,00
std.dev.	2,68			
numbers removed	0	0	0	0
stoplist tokens removed	0	0	0	0
stoplist types removed	0	0	0	0
1-letter words	0	0	0	0
2-letter words	0	0	0	0
3-letter words	0	0	0	0
4-letter words	0	0	0	0
5-letter words	19	1	3	0

Figura 10. Ejemplo de la aplicación *Wordlist*.

La información cuantitativa obtenida a través de la aplicación *Wordlist* se podía guardar en hojas de cálculo del programa *Excel*<sup>39</sup> que luego se utilizaban para el procesamiento estadístico con el programa SPSS, que se explica a continuación.

#### 9.7.2. Procesamiento estadístico (SPSS)

Tras las fases de etiquetado y cuantificación de los datos, como queda descrito en los apartados anteriores, se ha realizado el procesamiento estadístico. Para ello se ha utilizado la herramienta informática de análisis estadísticos SPSS<sup>40</sup> versión 13.0.

Siguiendo la orientación de investigadores como Lazaraton (2005) y Ortega y Iberri-Shea (2005), en esta tesis se ha intentado no limitarse solamente a la estadística descriptiva, sino ir más allá de la descripción estadística. Para ello, hemos recurrido a la estadística inferencial que nos va a permitir profundizar en el estudio de la muestra, tal y como lo hicieron Martínez (2005), Laufer y Girsai (2008), Gallego y Llach (2009) en sus investigaciones.

En primer lugar, era necesario hacer las pruebas de estadística descriptiva para, como su nombre indica, conocer mejor la muestra y ofrecer una visión general del conjunto de datos obtenidos. Hay tres tipos de pruebas de la estadística descriptiva que se usan comúnmente en el campo de la ASL: medidas de frecuencia, medidas de tendencia central y medidas de variabilidad o dispersión (Mackey y Gass, 2005). A partir de la obtención de esos resultados, se podía dar paso a la comprobación de diferencia significativa entre los niveles.

En segundo lugar, para comparar los resultados obtenidos entre los bloques para cada una de las variables y verificar si había diferencia estadísticamente significativa, se tenía que recurrir a la estadística inferencial. Ese conjunto de técnicas permite conocer si los resultados obtenidos pueden generalizarse a una población más amplia. Como casi siempre es muy difícil conseguir datos del total de individuos de una población, se suele reunir datos de una parte de ella y luego, con la ayuda de la estadística inferencial, generalizar los

---

<sup>39</sup> El *Excel* es una herramienta informática que trabaja con hojas de cálculo de Microsoft Office (2007).

<sup>40</sup> *Statistical Package for the Social Sciences* (SPSS) es un programa estadístico informático muy usado en las ciencias sociales creado en 1968 por Norman H. Nie, C. Hadlai Hull y Dale H. Bent (Stevens, 2002).

resultados a la totalidad de la población (Mackey y Gass, 2005). Entre los resultados de esas pruebas estadísticas están las medias, medianas, valores mínimos y máximos, y datos de desviación estándar. Entre las técnicas más aplicadas en estadística inferencial hay: ANOVA, prueba de chi cuadrado, análisis factorial, correlación y regresión lineal.

La prueba más utilizada en estadística inferencial es la ANOVA<sup>41</sup>, una prueba de tipo paramétrico<sup>42</sup>. Investigadores como Woodall (2002), Jegerski *et al.* (2011), Lu (2011), Salsbury *et al.* (2011) han utilizado ANOVA en sus estudios sobre la adquisición de L2. Sin embargo, el uso correcto de la estadística inferencial tipo ANOVA requiere que se cumplan algunos supuestos teóricos, ya que al no cumplirse alguno de ellos, se incurre en errores que comprometen los resultados obtenidos (Mackey y Gass, *op. cit.*).

Uno de los más importantes de estos supuestos es el de la normalidad<sup>43</sup> de la muestra. Para comprobar si los datos presentan una distribución normal, se recomienda la utilización de la prueba de Shapiro-Wilkinson en muestras con tamaño inferior a 2.000 y la de Kolmogorov-Smirnov para muestras superiores a 2.000 elementos. Otro de los supuestos de las pruebas ANOVA es la homogeneidad de las varianzas, también conocida como homocedasticidad. Para comprobar esta condición, se suele emplear un test tipo Levène, una prueba considerada más robusta para ese fin (Stevens, 2002).

La no confirmación de que las condiciones para la realización de estadística inferencial tipo ANOVA (es decir, normalidad y homocedasticidad de las variables estudiadas) no imposibilita la realización de pruebas de inferencia. Como dice Stevens (*op.cit.*), cuando los

---

<sup>41</sup> El análisis de la varianza (ANOVA) es una colección de modelos estadísticos que fueron inicialmente desarrollados por el estadístico y genetista R. A. Fisher. El análisis de varianza se utiliza para verificar si hay diferencias estadísticamente significativas entre medias cuando hay necesidad de comparar más de dos grupos. Los resultados del test ANOVA producen un valor F, que es la ratio de la variación entre los grupos sobre la variación dentro de los grupos (Mackey y Gass, 2005). Una varianza grande indica que hay mucha variación entre los sujetos, que hay mayores diferencias individuales con respecto a la media; una varianza pequeña nos indica poca variabilidad entre los sujetos, diferencias menores entre los sujetos (Vallejo, 2008).

<sup>42</sup> Los procedimientos paramétricos se refieren a pruebas estadísticas específicas para el análisis de datos que tienen una distribución normal (Mendenhall, 2007). Son pruebas paramétricas: la prueba t, ANOVA, la prueba de correlación y la de Pearson. Ejemplos de pruebas no paramétricas son chi cuadrado, la correlación de Spearman, la prueba U de Mann-Whitney, la W de Wilcoxon y la de Kruskal-Wallis.

<sup>43</sup> En estadística, la distribución normal describe los agrupamientos de los comportamientos o puntuaciones en una gráfica que tiene una curva en forma de campana y es simétrica según un determinado parámetro. Esta curva se conoce como campana de Gauss. En una distribución normal, las medidas de tendencia central (media, moda y mediana) coinciden en el punto central (Mackey y Gass, 2005).

supuestos de normalidad y homocedasticidad no se cumplen, se sugiere la utilización de pruebas de tipo no-paramétrico, que en lugar de la media, utilizan la mediana para calcular la significación (el valor-p).

Tras la realización de la prueba de normalidad (Shapiro-Wilk), se ha verificado que los resultados obtenidos indicaban que los datos de la muestra de la presente investigación presentaban valores de significación menores que 0,05 (valor- $p < 0,05$ ), lo que indicaba que se podía rechazar la hipótesis de que los datos cumplían con el supuesto de normalidad. Por ello, en lugar de utilizar pruebas paramétricas como ANOVA, se ha tenido que recurrir a pruebas de tipo no paramétrico equivalentes para verificar si había diferencia estadísticamente significativa entre los niveles para cada una de las variables analizadas. Así, para responder las preguntas de investigación planteadas en esta tesis, se han realizado pruebas no-paramétricas como la de Kruskal-Wallis<sup>44</sup> (ej.: Farahani y Meraji, 2011).

Una vez obtenidos los resultados de las pruebas no-paramétricas de Kruskal-Wallis, se ha identificado la necesidad de averiguar entre qué bloques se encontraba la diferencia estadísticamente significativa. Eso es así porque esa prueba sólo indica si hay diferencia entre los bloques, pero no identifica entre cuáles de ellos se encuentra la diferencia. Consecuentemente, hemos recurrido a pruebas tipo *post-hoc* como el test de Games-Howell, que sirve para identificar entre qué niveles había diferencia estadísticamente significativa para cada una de las variables estudiadas (fluidez y diversidad léxica, errores léxicos y funcionales, errores según la clase de palabra léxica, tipo de error, origen del error, tipo y origen del error).

---

<sup>44</sup> El test estadístico conocido como Kruskal-Wallis sirve para comparar tres o más grupos independientes y es la opción que se utiliza cuando no se cumplen los supuestos teóricos requeridos para la realización de un test ANOVA (Mackey y Gass, 2005).

## PARTE IV: RESULTADOS, ANÁLISIS Y DISCUSIÓN

### 10. RESULTADOS Y ANÁLISIS

El objetivo principal de la parte IV es presentar, analizar y discutir los resultados obtenidos a partir del *corpus* lingüístico recopilado en las 365 redacciones escritas en español por 142 estudiantes brasileños en formación como profesores de ELE. Además, se analizarán los datos del cuestionario que se aplicó a los participantes para conocer la percepción que tienen de la proximidad lingüística entre el español y el portugués.

Antes de abordar la exposición de resultados y su análisis, explicamos el proceso de codificación de los datos lingüísticos que fue necesario realizar previamente para poder identificar y clasificar los errores (véase apartado 9.7. 'Instrumentos para el procesamiento y análisis de los datos lingüísticos'). La codificación de los datos se realizó en cinco fases (tabla 5).

Tabla 5. Fases de codificación, códigos y número de palabras etiquetadas.

Fases de codificación	Clasificación	Código	N palabras
Fase 1	Palabras total		43.476
	Palabras léxicas	<CW>	23.935
	Palabras funcionales	<FW>	19.541
Fase 2	Errores	<ERR>	4.407
	Errores – palabras léxicas	<CW><ERR>	3.302
	Errores – palabras funcionales	<FW><ERR>	1.105
Fase 3a	Clase de palabra léxica		23.935
	Sustantivos	<CW><SUS>	10.283
	Adjetivos	<CW><ADJ>	3.237
	Verbos	<CW><VER>	7.184
Fase 3b	Adverbios	<CW><ADV>	3.231
	Clase de palabra léxica con error		3.302
	Sustantivos	<CW><SUS><ERR>	1.153
	Adjetivos	<CW><ADJ><ERR>	497
	Verbos	<CW><VER><ERR>	1.359
Fase 4	Adverbios	<CW><ADV><ERR>	293
	Tipo de error léxico		3.302
	Errores formales	<F>	2.427
	Errores semánticos	<C>	467
Fase 5	Errores mixtos	<M>	408
	Origen del error léxico		3.302
	Errores interlingüísticos	<INTER>	1.644
	Errores intralingüísticos	<INTRA>	1.487
	Errores ambiguos	<AMB>	171

En la fase 1 se etiquetaron todas las palabras del *corpus* lingüístico (N=43.476) en palabras léxicas (N=23.935: sustantivos, adjetivos, verbos, adverbios) y palabras funcionales (N=19.541: artículos, pronombres, preposiciones, conjunciones, verbos auxiliares, etcétera).

En la fase 2 se etiquetaron los errores, tanto en las palabras léxicas (N=3.302) como en las palabras funcionales (N=1.105).

En la fase 3a, centramos la atención en las palabras léxicas, y las etiquetamos según la clase: sustantivo (N=10.203), adjetivo (N=3.237), verbo (N=7.184) y adverbio (N=3.231). En la fase 3b, identificamos las clases de palabras léxicas con errores: sustantivo (N=1.153), adjetivo (N=497), verbo (N=1.359) y adverbio (N=293).

En la fase 4, se etiquetaron los errores según el tipo, es decir, si el error afectaba a la forma (N=2.427), al contenido (N=467) o a ambos aspectos (errores mixtos: N=408) de la palabra léxica.

En la quinta y última fase del proceso de etiquetado de los datos, centramos la atención en el origen del error de las palabras léxicas. Así, se codificaron los errores en interlingüísticos (N=1.644), si el error se debía a las características de la L1 del participante, intralingüísticos (N=1.487), si el error se debía al propio proceso de la interlengua, y ambiguos (N=171), cuando no se podía identificar con exactitud el origen del error.

Con el fin de poder observar y analizar los posibles cambios que se dan a lo largo del período de formación, agrupamos las producciones de los participantes en tres bloques de análisis siguiendo los criterios descritos en el apartado 9.6. (Organización del *corpus* lingüístico).

Para el análisis cuantitativo de los datos realizamos pruebas estadísticas. Como se explicó anteriormente en el apartado 9.7.2. Procesamiento estadístico (SPSS), el test estadístico ANOVA es la prueba más conocida y utilizada para comparar medias. Sin embargo, para que se utilice el test ANOVA, un test de tipo paramétrico, los datos tienen que cumplir con el supuesto de normalidad. Dado que los datos de esta investigación no presentan una distribución normal (véase anexo 6, tabla 32) recurrimos a un test de tipo no paramétrico como el de Kruskal-Wallis. Esta prueba estadística utiliza los valores de la mediana, en lugar de las medias, para calcular la diferencia entre los bloques. Sin embargo, este test no indica exactamente entre cuáles de los tres bloques de análisis se encuentra la diferencia



observada. Por lo tanto, tras realizar la prueba de Kruskal-Wallis, el siguiente paso era recurrir a un test estadístico también de tipo no paramétrico, como el test U de Mann-Whitney, que nos permitiera identificar entre qué bloques se observaba diferencia estadísticamente significativa, es decir, si la diferencia se encontraba entre los bloques 1 y 2, los bloques 2 y 3, o solamente entre los bloques 1 y 3. Con el objetivo de contestar las preguntas de investigación planteadas previamente, se siguió este procedimiento en el análisis de cada una de las variables como se verá a continuación. En los apartados siguientes presentamos los resultados obtenidos en el proceso de análisis de los datos con el fin de responder a cada una de las preguntas de investigación así como ejemplos extraídos del *corpus* para ilustrar los errores que se han identificado.

### 10.1. Resultados y análisis de la fluidez y de la diversidad léxica

**Pregunta 1: En la producción escrita de los profesores brasileños de ELE en formación ¿cambia la fluidez y la diversidad léxica a lo largo del período de formación?**

Para contestar la primera pregunta que planteamos en esta investigación, realizamos un análisis global del *corpus* lingüístico para detectar el grado de fluidez y diversidad léxica en cada uno de los tres bloques de análisis. Para ello, recurrimos a los datos de la fase 1 de codificación. El número total de palabras (N=43.476), que constituye el *corpus* lingüístico se ha clasificado en palabras léxicas (N=23.935; 55,05%) y palabras funcionales (N=19.541; 44,95%). A continuación presentamos los resultados de la aplicación de la medida de fluidez que consiste en el cómputo del número total de palabras por texto (Wolfe-Quintero, *et al.*, 1998; Jarvis *et al.*, 2003).

Tal como se presenta gráficamente en la figura 11, la media del número total de palabras (*tokens* total) identificadas en los textos del bloque 1 (M=101) es inferior a la media de palabras encontradas en las producciones del bloque 2 (M=122,9) y a su vez, esta media es inferior a la del bloque 3 (M=132,6). De manera general, se observa un aumento de la **fluidez** a medida que aumenta el número de horas de instrucción  $Bq1 < Bq2 < Bq3$ . Para ver si el aumento de la fluidez es estadísticamente significativo, aplicamos pruebas estadísticas que explicamos a continuación.

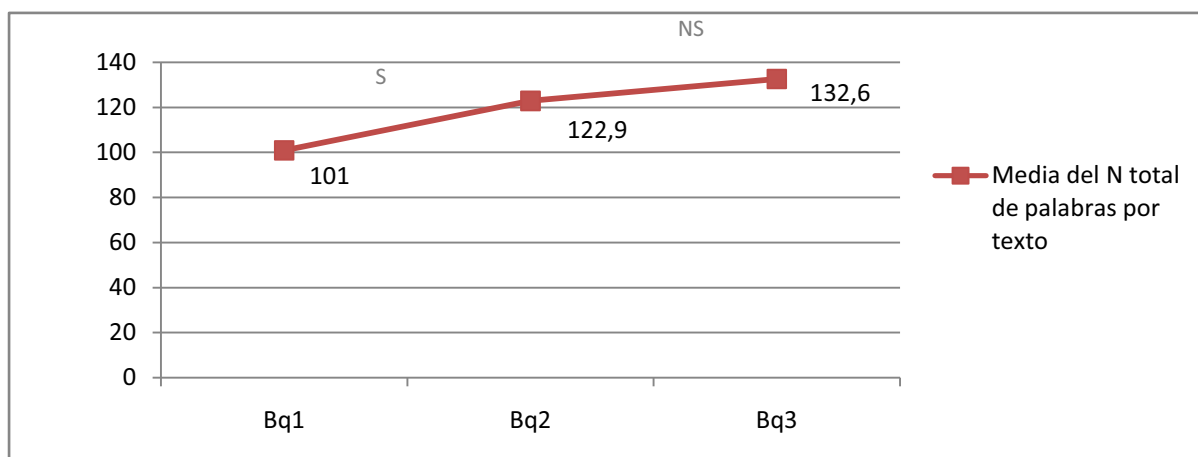


Figura 11. Medida de fluidez.

Nota: S=diferencia significativa; NS=diferencia no significativa.

Como hemos explicado en la introducción de este capítulo, los datos de la presente investigación no cumplen con el presupuesto de normalidad, por lo que se ha tenido que recurrir a pruebas de tipo no paramétrico para comprobar la significación de la diferencia observada entre los tres bloques de análisis. La primera prueba utilizada es el test de Kruskal Wallis, cuyos resultados para la variable de fluidez indican que hay diferencia estadísticamente significativa entre por lo menos dos de los tres bloques (valor- $p=0,000$ , véase anexo 7, tabla 39). En la tabla 6 abajo vemos los resultados del test U de Mann-Whitney, que revela que la diferencia entre las medias del número total de palabras por texto que se había observado inicialmente tanto entre los bloques de análisis 1 y 2 ( $Bq1 < Bq2$ ) como entre los bloques 1 y 3 ( $Bq1 < Bq3$ ) es estadísticamente significativa (valor- $p=0,000$ ). Sin embargo, aunque hayamos observado que entre el bloque 2 y 3 había diferencia entre las medias, ésta no es estadísticamente significativa (valor- $p=0,087$ ).

Tabla 6. Test U de Mann-Whitney: variable fluidez.

Variable fluidez: N total de palabras				
Bloques	N	Rango medio	Z	Valor-p
1	123	97,51	-4,488	<b>,000</b>
2	108	137,06		
2	108	112,94	-1,709	,087
3	134	128,40		
1	123	101,70	-5,641	<b>,000</b>
3	134	154,06		

Por lo tanto podemos afirmar que hay un aumento significativo de la **fluidez**, es decir, los participantes son capaces de escribir textos más largos después de alcanzar el nivel intermedio (bloque 2). Sin embargo, a pesar de que entre el nivel intermedio y avanzado haya una tendencia al aumento de la fluidez, ésta no es estadísticamente significativa. Ello podría explicarse porque han dispuesto del mismo tiempo para la realización de la tarea (redacciones) que ha sido constante a lo largo de los tres niveles. Sin embargo, cabe ver si la calidad de los textos ha aumentado porque presentan una mayor riqueza léxica.

Para explorar esta cuestión, hemos aplicado la medida de diversidad léxica Uber U (Dugast, 1978, 1979, 1980), que también utiliza el número de *tokens* y de *types* presentes en cada texto y que se verá a continuación. La decisión de aplicar la medida Uber U en lugar de la tradicional ratio del número de *types* dividido por el número de *tokens* (RTT) se debe a la necesidad de evitar el efecto del tamaño del texto que suele ocurrir con la RTT y que distorsiona los resultados cuando los textos analizados no tienen el mismo número de palabras (Wolfe Quintero *et al.*, 1998; Vermeer, 2000; Jarvis, 2002; Dewaele y Pavlenko, 2003). Asimismo, dado que nuestro *corpus* está formado por textos comprendidos entre 24 y 235 palabras, el índice Uber U es la medida más indicada para medir la diversidad léxica en textos pequeños (Vermeer, *op. cit.*; Jarvis, *op. cit.*).

El análisis de las medias de los bloques para la variable **diversidad léxica** revela un comportamiento no esperado, porque los datos indican que los textos del bloque 1 (M=26,91) presentan una media del índice de diversidad superior a la de los textos del bloque 2 (M=25,16) y 3 (M=26,43). Además, los textos del bloque 2 presentan una media del índice de diversidad inferior a los bloques 1 y 3. Finalmente, el bloque 3 presenta una media del índice de diversidad superior al bloque 2, pero inferior al bloque 1 ( $Bq1 > Bq2 < Bq3$ ) (figura 12).

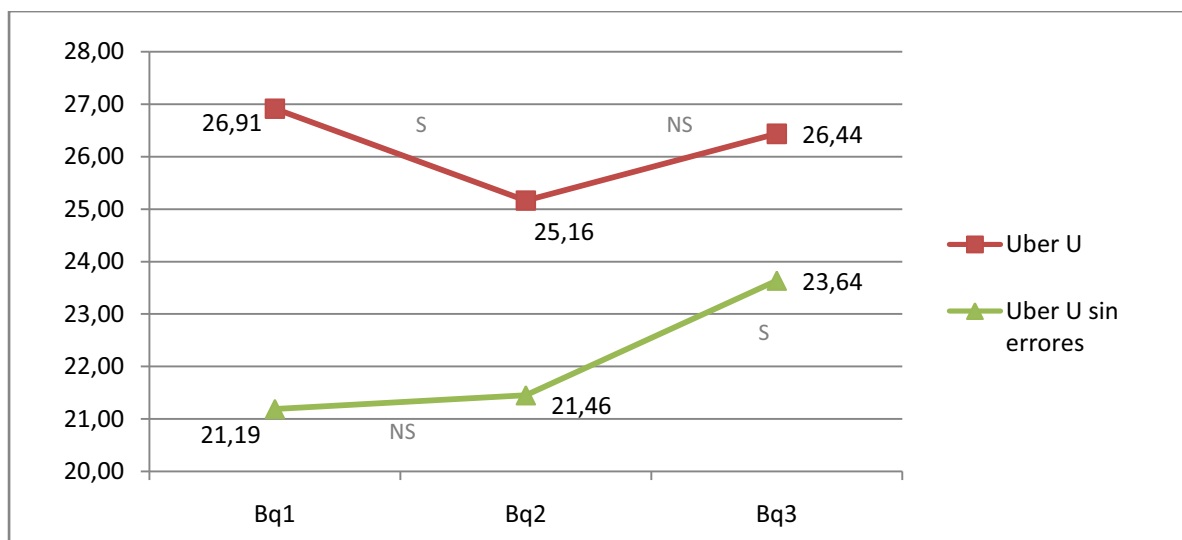


Figura 12. Medida de diversidad léxica: Uber U y Uber U sin errores.  
 Nota: S=diferencia significativa; NS=diferencia no significativa.

Este resultado nos llevó a analizar cualitativamente los datos y averiguamos que el número elevado de errores léxicos en los textos del bloque 1 (véase 10.2. Resultados y análisis de errores léxicos y funcionales) era la razón por la que se observa un alto índice de diversidad léxica. Un mayor número de errores ortográficos representa más variedad, ya que, por ejemplo, si en un texto encontramos para la palabra española ‘conocí’ (conocí, esp./conheci, port.) formas como ‘conocí’, ‘conoci\*’ y ‘coneci\*’, se computan tres *types*. Ello nos ha llevado, siguiendo a Engber (1995), a aplicar la prueba de diversidad léxica excluyendo los errores tal como sugiere en su propio estudio<sup>45</sup>.

Como vemos en la figura 12, las medias de diversidad para la medida Uber U sin errores son más bajas, pero aumentan progresivamente a lo largo de los tres niveles. Esto refleja un perfil de desarrollo más claro de la diversidad léxica, sin que ésta quede enmascarada por los errores. Así, observamos que la media del bloque 1 (M=21,19) es inferior a la del bloque 2 (M=21,46) y ésta es a su vez inferior a la del bloque 3 (M=23,64) (Bq1<Bq2<Bq3).

Tras la comprobación de que los datos no cumplían con el presupuesto de normalidad (anexo 6, tabla 32), aplicamos la prueba de Kruskal-Wallis, que reveló diferencia

<sup>45</sup> Engber (1995) estudió la variedad o diversidad léxica a través de la medida de variedad léxica libre de errores (*error-free lexical variation measure*) que consistía en la suma de los diferentes ítems léxicos por segmento menos la suma del número total de errores léxicos por segmento dividido por la suma del número total de ítems léxicos por segmento.

estadísticamente significativa entre las medias tanto para la variable Uber U (valor-p=0,05) como para Uber U sin errores (valor-p=0,000) (véase anexo 7, tabla 40).

En la tabla 7 presentamos los resultados de la prueba no paramétrica de Mann-Whitney para el índice Uber U, que incluye también las palabras léxicas con error. Estos resultados indican que entre los bloques 1 y 2 hay diferencia estadísticamente significativa (valor-p=0,023), mientras que entre los bloques 2 y 3 (valor-p=0,064) y también entre los bloques 1 y 3 el aumento no es significativo (valor-p=0,577).

Tabla 7. Test U de Mann-Whitney: variable diversidad léxica (Uber U)

<b>Variable: Diversidad léxica (Uber U)</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	125,40	-2,281	<b>,023</b>
2	108	105,30		
2	108	112,21	-1,854	,064
3	134	128,99		
1	123	131,70	-,558	,577
3	134	126,52		

Sin embargo, nos interesa analizar los datos de la tabla 8 que contiene la información de esta misma prueba para el índice 'Uber U sin errores' que elegimos utilizar para medir la diversidad léxica.

Tabla 8. Test U de Mann-Whitney: variable diversidad léxica (Uber U sin errores).

<b>Variable: Diversidad léxica (Uber U sin errores)</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	114,34	-,403	,687
2	108	117,89		
2	108	103,22	-3,646	<b>,000</b>
3	134	136,23		
1	123	110,48	-3,828	<b>,000</b>
3	134	146,00		

La prueba U de Mann-Whitney para la variable **diversidad léxica** (Uber U sin errores) revela que entre los niveles inicial (bloque 1) e intermedio (bloque 2) no hay diferencia estadísticamente significativa (valor-p=0,687). Por otro lado, sí hay diferencia estadísticamente significativa entre los niveles intermedio (bloque 2) y avanzado (bloque 3)

y también entre los niveles inicial (bloque 1) y avanzado (bloque 3) (valor- $p=0,000$ ). Este resultado indica que la diversidad léxica aumenta a medida que aumenta el número de horas de instrucción, pero esta mejora sólo se verifica después que los alumnos alcanzan el nivel avanzado.

En resumen, la **fluidez** aumenta entre los niveles inicial e intermedio, lo que revela que la producción de los estudiantes presenta más palabras léxicas a partir del nivel intermedio. Aunque se haya visto un aumento continuado en el uso de las palabras léxicas en el nivel avanzado, éste no es estadísticamente significativo. Por otro lado, el incremento que se ha observado en la **diversidad léxica** sólo es significativo entre los niveles intermedio y avanzado, es decir, la producción escrita de los estudiantes presenta más variedad en el uso de vocabulario después de que éstos alcanzan el nivel avanzado. Sin embargo, es importante tener en cuenta que en la medida de fluidez se han computado todas las palabras léxicas, con y sin errores, mientras que en el análisis de la diversidad se ha optado por excluir los errores léxicos. Así, ambos resultados quedan afectados por la inclusión o exclusión del error léxico.

## 10.2. Resultados y análisis de errores léxicos y funcionales

**Pregunta 2: ¿En qué categoría de palabra (léxica o funcional) se observa un mayor número de errores según el número de horas de instrucción?**

En este apartado centramos nuestra atención en los errores identificados ( $N=4.407$ ) en el *corpus* lingüístico durante la fase 2 de codificación (tabla 5). Por lo que respecta a los cambios observados en cada uno de los bloques, en la tabla 9 presentamos la media de errores identificados por texto en cada uno de los niveles. Estos datos nos permiten constatar que la media del número total de errores (*tokens*) del bloque de análisis 1 ( $M=14,6$ ) es mayor que en el bloque de análisis 2 ( $M=12,6$ ), y que éstas son superiores a la media de errores identificados en los textos del bloque 3 ( $M=9,2$ ) ( $Bq1>Bq2>Bq3$ ).

Tabla 9. Media de errores por texto.

Bloques	Media del número total de errores por texto	
	<i>types</i>	<i>tokens</i>
Bq1	13,6	14,6
Bq2	11,8	12,6
Bq3	8,8	9,2

También se computaron los errores en las palabras únicas (*types*). Como se ha verificado una gran variedad de definiciones sobre qué se considera un *type* en las investigaciones que analizamos para desarrollar esta tesis, creemos oportuno comentar algunas de ellas. Generalmente, con el objetivo de examinar la diversidad lexical como producto, se consideran cada una de las flexiones de una misma palabra como un *type* (ej.: Laufer, 1991; Engber, 1995; Yu, 2007), mientras que otras investigaciones distinguen entre los tipos de flexión que presenta mismo lexema con el objetivo de conocer más sobre el desarrollo morfológico de los sujetos participantes (ej.: Durán *et al.*, 2004). En la presente investigación consideramos cualquier diferencia gráfica entre una palabra y otra, aunque sea solo una tilde o una letra, como un nuevo *type*. Esto es así porque el programa informático utilizado *Wordsmith Tools 5* (Scott, 2010) identifica cada diferencia como un nuevo ítem y procesa los datos de esta manera. Además, nos interesaba conocer las variaciones ortográficas que se dan de un mismo referente en nuestra base de datos. Creemos que esto es especialmente relevante en el contexto de formación de profesores y de lenguas próximas como es nuestro caso para que se pueda profundizar en el análisis del tipo y origen del error léxico. Así, en la tabla 9 vemos que se verifica también una disminución de la media de los errores por texto en las palabras únicas (*types*) a lo largo de los tres niveles (Bq1>Bq2>Bq3). La comparación de la media de *types* con la de *tokens* revela que hay poca diferencia entre ellas, lo que manifiesta de forma clara que existe gran variedad de errores en cada uno de los tres niveles, por lo que, de ahora en adelante, se considerarán los *tokens* en el análisis de los errores.

Del número total de errores en el *corpus* (N=4.407), se identificaron **errores léxicos** (N=3.302; 75%) y **errores funcionales** (N=1.105; 25%) (tabla 5). El hecho de que los errores identificados en las palabras léxicas sean los más abundantes y los que más entorpecen la

comunicación entre un hablante de L1 y un hablante de L2 (Fernández, 1997) son algunos de los argumentos que ayudan a justificar el interés y la importancia del presente estudio. En la tabla 10 presentamos algunos ejemplos de los errores léxicos y funcionales encontrados en el *corpus* analizado en la presente investigación.

Tabla 10. Ejemplos de errores léxicos y funcionales.

léxicos		funcionales	
error (interlengua)	español (L2)	error (interlengua)	español (L2)
dificultad*	dificultad	apesar* de	a pesar de
vocábolos*	vocablos	me* relación	mi
(el aprendizaje del español) pasó a volverse* una necesidad	ser	pra*	para
(...) hay profesores acostumbrados a hacer* alumnos para el vestibular.	preparar	com*	con

En la figura 13 observamos gráficamente el comportamiento de las medias de estas dos variables a lo largo de los tres niveles.

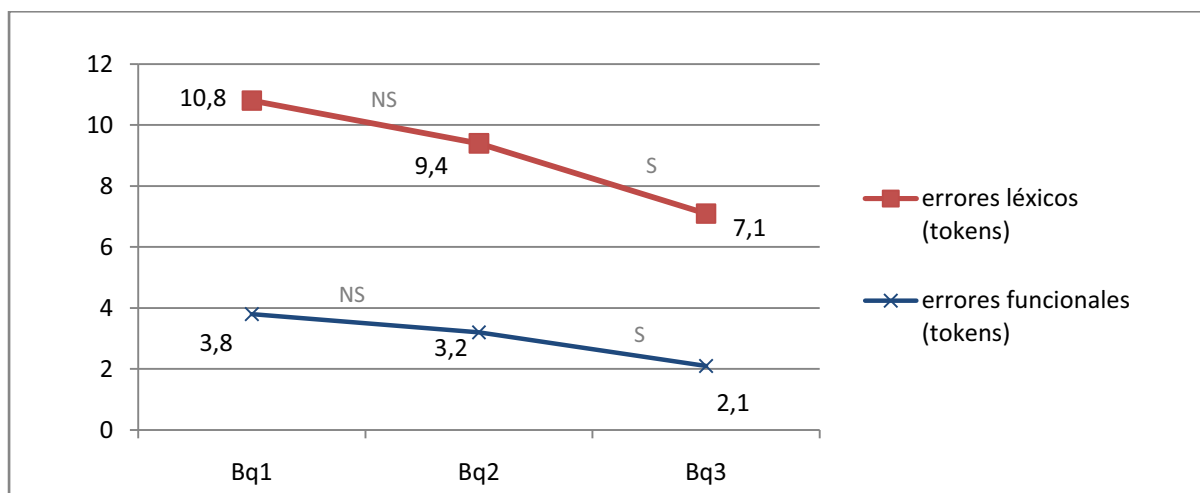


Figura 13. Media del número errores léxicos y funcionales por texto.

Nota: S=diferencia significativa; NS= diferencia no significativa.

Con respecto a la variable **errores léxicos**, verificamos que la media del bloque de análisis 1 (M=10,8) es superior a la media del bloque 2 (M=9,4) y ambas son superiores a la media del bloque 3 (M=7,2) (Bq1>Bq2>Bq3). Constatamos que en el caso de la variable **errores**



**funcionales** ocurre lo mismo ( $Bq1 > Bq2 > Bq3$ ), es decir, la media de errores identificados en las palabras funcionales en los textos del bloque 3 ( $M=2,1$ ) es inferior a los de los bloques 2 ( $M=3,2$ ), y éstos, a su vez, presentan una media inferior a la media del bloque 1 ( $M=3,8$ ).

En la tabla 33 (anexo 6) se puede consultar el resultado del test de normalidad para las variables errores léxicos y funcionales. Los resultados del test no paramétrico de Kruskal-Wallis indican que hay diferencia estadísticamente significativa entre los bloques (valor- $p=0,000$ , anexo 7, tabla 41). Sin embargo, como esta prueba no especifica entre cuáles de los tres niveles se encuentra la diferencia observada, recurrimos a la prueba U de Mann-Whitney para conocer entre qué bloques de análisis hay diferencia estadísticamente significativa.

En la tabla 11 se puede observar que, para las variables errores funcionales (valor- $p=0,283$ ) y errores léxicos (valor- $p=0,081$ ), la diferencia observada inicialmente entre los bloques 1 y 2 ( $Bq1 > Bq2$ ) no es estadísticamente significativa.

Tabla 11. Test U de Mann-Whitney: variable errores léxicos y funcionales.

<b>Variable: errores funcionales (tokens)</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	120,38	-1,073	,283
2	108	111,01		
2	108	139,99	-3,738	,000
3	134	106,60		
1	123	152,28	-4,868	,000
3	134	107,63		
<b>Variable: errores léxicos (tokens)</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	123,18	-1,746	,081
2	108	107,82		
2	108	138,91	-3,481	,001
3	134	107,47		
1	123	156,84	-5,734	,000
3	134	103,45		

Por otro lado, se puede observar que tanto en el caso de los bloques 2 y 3 ( $Bq2 > Bq3$ ) como en el de los bloques 1 y 3 ( $Bq1 > Bq3$ ) hay diferencia estadísticamente significativa (valor- $p = 0,000$ ).

En resumen, comprobamos que el número de **errores léxicos** es superior al número de **errores funcionales**. Además, entre los niveles intermedio y avanzado tanto en el caso de los errores léxicos como en el de los funcionales la media de errores presentes por texto disminuye según aumenta el número de horas de instrucción. Sin embargo, es importante resaltar que la disminución sólo es estadísticamente significativa cuando los alumnos alcanzan el nivel avanzado, lo que indica que es a partir de este nivel que los estudiantes empiezan a tener un uso más correcto del léxico que conocen.

### 10.3. Resultados y análisis de los errores según la clase de palabra léxica

**Pregunta 3: ¿A qué clase de palabra léxica (nombre, adjetivo, verbo o adverbio) afectan los errores a medida que aumenta el número de horas de instrucción?**

En este apartado centramos nuestra atención en las palabras léxicas (sustantivos, adjetivos, verbos y adverbios) y en los errores identificados en este grupo de palabras en la fase 3 de codificación. Con el objetivo de poder contestar la tercera pregunta de esta investigación, clasificamos todas las palabras léxicas según la clase: sustantivos, adjetivos, verbos y adverbios. Esta clasificación nos permite identificar el uso que hacen los participantes de las palabras léxicas y en qué clase se observan más errores.

Primeramente, analizamos los resultados de la fase 3a de codificación de los datos que nos revela que el número total de sustantivos encontrados en el *corpus* lingüístico (SUS: N=10.283) es superior a las demás clases de palabras léxicas. La segunda clase más abundante es la de los verbos (VER: N=7.184), seguida de los adjetivos (ADJ: N=3.237) y de los adverbios (ADV: N=3.231). El número total de adjetivos y adverbios presentes en el *corpus* lingüístico es bastante similar.

Posteriormente, analizamos los resultados de la fase 3b de codificación, en la que vimos que del total de errores léxicos identificados (N=3.302) la mayoría de ellos afectaba a los verbos (ERR\_VER: N=1.359) y a los sustantivos (ERR\_SUS: N=1.153). Los adjetivos (ERR\_ADJ: N=497) y los adverbios (ERR\_ADV: N=293) por el contrario, fueron las categorías léxicas en las que se encontraron menos errores.

En la tabla 12 presentamos las medias de palabras léxicas (*tokens*) por texto según la clase (sustantivo, adjetivo, verbo o adverbio) y la media de errores por texto según la clase en cada bloque de análisis.

Tabla 12. Media de palabras léxicas y media de errores en las palabras léxicas por texto según la clase.

Bloques	sustantivo		adjetivo		verbo		adverbio	
	total	error	total	error	total	error	total	error
Bq1	23,5	3,5	7,4	1,47	16,9	4,8	7,5	1
Bq2	29	3	9,5	1,52	20,2	4	9	0,9
Bq3	31,8	2,9	9,7	1,13	21,8	2,5	9,9	0,6

En la figura 14 vemos gráficamente el comportamiento de estas variables a lo largo de los tres niveles.

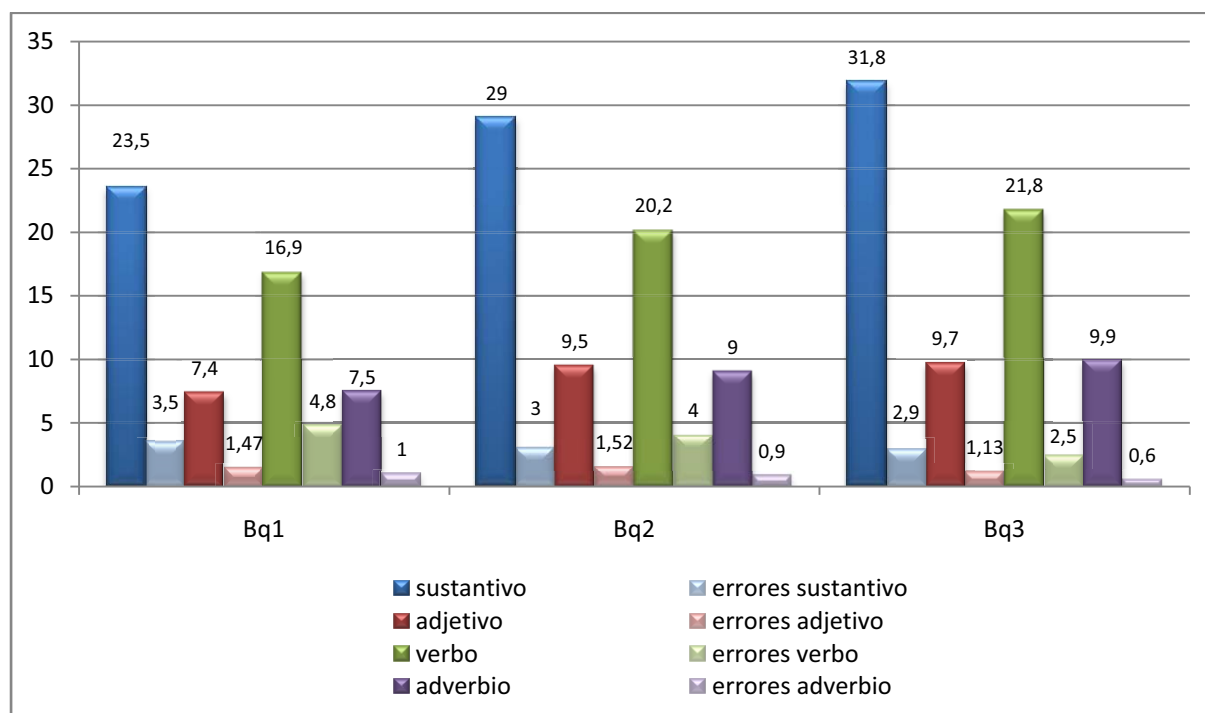


Figura 14. Media de palabras léxicas y media de errores en las palabras léxicas por texto según la clase.

Estos datos nos permiten comprobar que para todas las variables de la fase 3a de codificación (sustantivos, adjetivos, verbos y adverbios) la media de palabras del bloque de análisis 1 es menor que en el bloque 2, y que éstas son inferiores a la media de palabras léxicas según la clase en los textos del bloque 3 ( $Bq1 < Bq2 < Bq3$ ), es decir, hay un aumento

en el uso de las palabras léxicas en todas las clases según aumenta el número de horas de instrucción. Asimismo, observamos también el comportamiento de las variables identificadas en la fase 3b de codificación (errores sustantivo, errores adjetivo, errores verbo, errores adverbio). En este caso ocurre lo contrario, es decir, las medias de errores por clase de palabra disminuyen a medida que aumenta el número de horas de instrucción (Bq1>Bq2>Bq3). En la tabla 13 presentamos algunos ejemplos de los errores identificados según la clase de palabra léxica.

Tabla 13. Ejemplos de errores léxicos según la clase de palabra.

Errores léxicos							
Sustantivo		Adjetivo		Verbo		Adverbio	
error (interlengua)	español (L2)	error (interlengua)	español (L2)	error (interlengua)	español (L2)	error (interlengua)	español (L2)
turma*	grupo	imediata*	inmediata	leia*	lea	facilmente*	fácilmente
vontade*	ganas	felizes*	felices	(...) traté de arreglar* un empleo.	conseguir	sempre*	siempre
bolsas*	becas	atenciosas*	atentas	(...) no sé si macharé* maestra	me haré	lá*	allá

Antes de empezar a analizar los datos correspondientes a los errores en las palabras léxicas para contestar la tercera pregunta planteada en esta investigación, creemos pertinente examinar el comportamiento de las variables referentes a las clases de palabras léxicas en cada uno de los tres niveles para conocer si el aumento en su uso se da de manera relevante a medida que los participantes avanzan en su proceso de aprendizaje. En la figura 15 se puede observar gráficamente el comportamiento de estas variables en cada uno de los tres bloques.

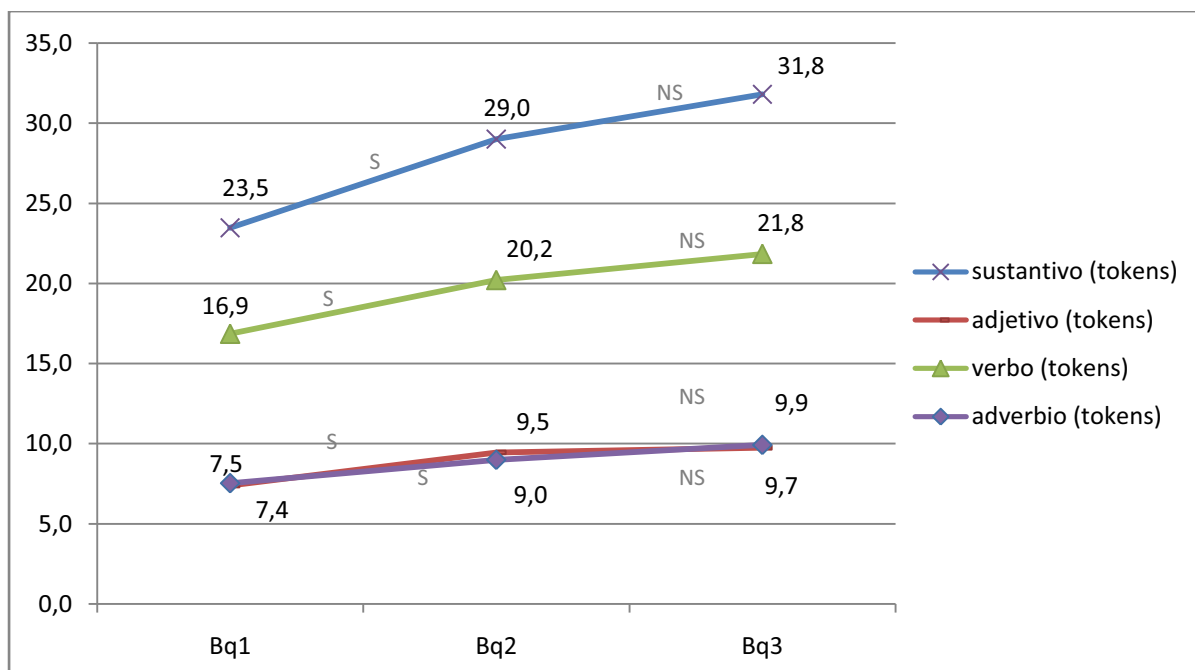


Figura 15. Media de palabras léxicas según la clase por texto.

Nota: S=diferencia significativa; NS=diferencia no significativa.

Con el fin de analizar estos datos, realizamos una serie de pruebas estadísticas. El test de normalidad Shapiro-Wilk para las variables clase de palabra léxica comprueba que los datos no siguen una distribución normal (anexo 6, tabla 34). Los resultados de la prueba Kruskal-Wallis indican que hay diferencia estadísticamente significativa entre los bloques para todas las variables (anexo 7, tabla 42). En la tabla 14 presentamos los resultados del test estadístico U de Mann-Whitney que utilizamos para conocer exactamente entre qué bloques de análisis se verifica la diferencia estadísticamente significativa.

En el caso de los **sustantivos** vemos que hay diferencia estadísticamente significativa entre los bloques de análisis 1 y 2 ( $Bq1 < Bq2$ ) así como entre los bloques 1 y 3 ( $Bq1 < Bq3$ ) (valor- $p=0,000$ ). Sin embargo, entre los bloques 2 y 3 ( $Bq2 < Bq3$ ) no se verifica diferencia estadísticamente significativa (valor- $p=0,063$ ). A partir de estos resultados podemos afirmar que los estudiantes del nivel intermedio emplean más sustantivos que los alumnos del nivel inicial en su producción escrita en español. Con todo, aunque la media de sustantivos identificados en el nivel avanzado era superior a la media del nivel intermedio, la diferencia entre ellas no es significativa estadísticamente. Por lo tanto, podemos decir que la cantidad

de sustantivos utilizados en las redacciones se mantiene estable entre los niveles intermedio y avanzado.

Tabla 14. Test U de Mann-Whitney: variable clases de palabra léxica.

<b>Variable: sustantivo</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	99,54	-3,999	<b>,000</b>
2	108	134,75		
2	108	112,19	-1,857	,063
3	134	129,00		
1	123	100,90	-5,810	<b>,000</b>
3	134	154,79		
<b>Variable: adjetivo</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	99,59	-3,997	<b>,000</b>
2	108	134,69		
2	108	120,50	-,199	,842
3	134	122,30		
1	123	108,24	-4,304	<b>,000</b>
3	134	148,06		
<b>Variable: verbo</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	101,06	-3,631	<b>,000</b>
2	108	133,02		
2	108	114,55	-1,388	,165
3	134	127,10		
1	123	106,02	-4,753	<b>,000</b>
3	134	150,10		
<b>Variable: adverbio</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	103,94	-2,938	<b>,003</b>
2	108	129,74		
2	108	115,08	-1,284	,199
3	134	126,67		
1	123	109,09	-4,126	<b>,000</b>
3	134	147,28		

A partir de los datos presentados en la tabla 14 vemos que el comportamiento observado para la variable sustantivos se repite en las demás variables, es decir, se verifica diferencia estadísticamente significativa entre los bloques 1 y 2 (valores- $p < 0,05$ ), pero no entre los bloques 2 y 3 (valores- $p > 0,05$ ). Pero el aumento en el uso de cada una de estas clases de palabras se ve reflejado en la diferencia estadísticamente significativa que se verifica entre los bloques 1 y 3 (valores- $p < 0,000$ ).

Este resultado nos permite afirmar que la producción de palabras léxicas aumenta de manera significativa para todas las clases de palabras léxicas entre los niveles inicial e intermedio (bloques 1 y 2). Sin embargo, en lo que se refiere al aumento de las medias de las palabras léxicas por texto entre los niveles intermedio y avanzado (bloques 2 y 3), las pruebas estadísticas indican que éste no es estadísticamente significativo para ninguna de las variables (valor- $p > 0,05$ ).

En resumen, el análisis de la producción escrita de la muestra revela que se da un aumento notable en el uso de las palabras léxicas (**sustantivo, adjetivo, verbo y adverbio**) en su producción escrita entre los niveles inicial e intermedio. Esto indica que los alumnos son capaces de producir una mayor variedad de vocabulario cuando alcanzan el nivel intermedio. Sin embargo, entre los niveles intermedio y avanzado, el aumento en la media de palabras léxicas por texto que se observa en cada una de las clases de palabra no es importante.

Como en la fase 2 de codificación ya habíamos identificado los errores en el *corpus* lingüístico, podemos ahora analizar también la ocurrencia que tienen en cada una de las clases de palabra léxica, con el objetivo de contestar a la tercera pregunta planteada en esta investigación. Con respecto a la variable **errores sustantivo**, verificamos que la media de errores presentes en el bloque de análisis 1 ( $M=3,52$ ) es superior a la media del bloque 2 ( $M=3,03$ ) y éstas son mayores que la media del bloque 3 ( $M=2,93$ ) ( $Bq1 > Bq2 > Bq3$ ).

En la figura 16 se puede apreciar más claramente estas tendencias y ver que aunque haya disminución en la media de errores a medida que aumenta el número de horas de instrucción entre los bloques de análisis 2 y 3, ésta es bastante sutil.

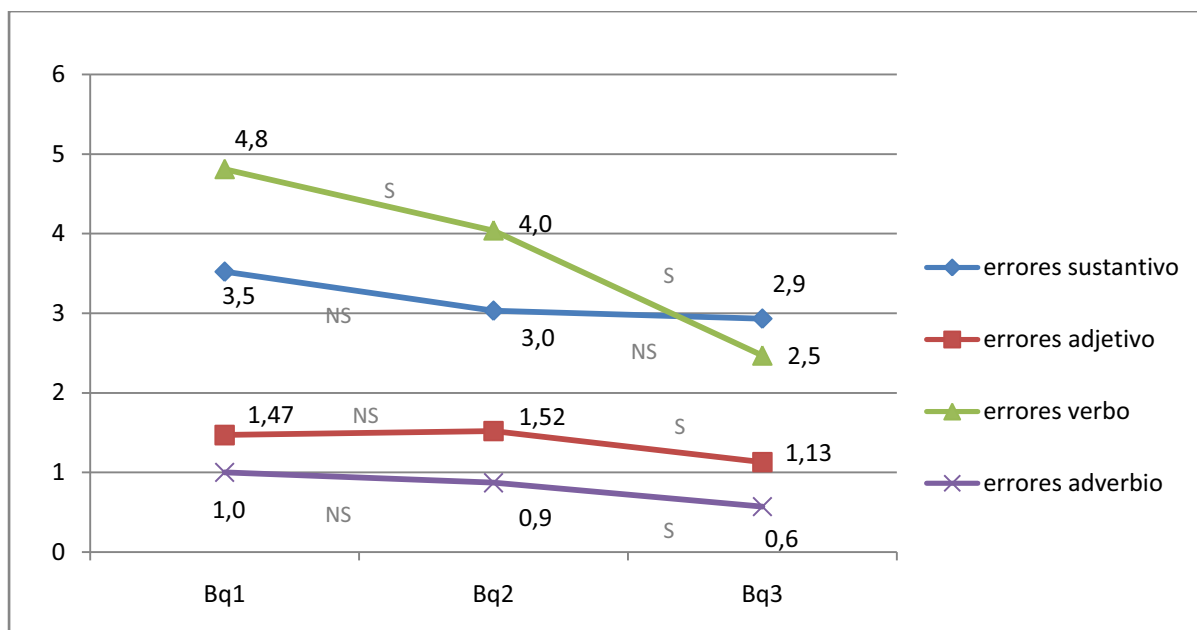


Figura 16. Media de errores léxicos según la clase.

Nota: S=diferencia significativa; NS=diferencia no significativa.

En lo que se refiere a la variable **errores adjetivo**, la media de errores tiene un comportamiento diferente, ya que no sigue la tendencia de disminuir a medida que aumenta el número de horas de instrucción que se ha observado en las demás variables. Aunque la media de errores encontrados en el bloque 3 (M=1,13) es inferior a la media del bloque 1 (M=1,47), la media de errores por texto del bloque 2 (M=1,52) es superior a las medias del bloque 1 y del bloque 3 (Bq1<Bq2>Bq3).

La media de errores de la variable **errores verbo** disminuye a medida que aumenta el número de horas de instrucción (Bq1>Bq2>Bq3). Pese a ser la categoría de palabra léxica con el mayor número de errores por texto tanto en el bloque 1 (M=4,81) como en el bloque 2 (M=4,04), es también la que presenta una disminución más visible de la media de errores entre los tres bloques, una vez que la media de errores en el bloque 3 es más baja (M=2,47). A simple vista podemos decir que el número de errores identificados en la clase de los verbos baja de forma acentuada a medida que aumenta el número de horas de instrucción. Igualmente, más adelante lo verificamos a través de la prueba estadística U de Mann-Whitney.



Por último, analizamos las medias referentes a la variable **errores adverbio**. El análisis de las medias indica que hay una disminución de este tipo de error a lo largo de los tres niveles. La media de errores identificados por texto es mayor en el bloque 1 (M=1,00) que la del bloque 2 (M=0,87) y que ésta es a su vez, superior a la media de errores del bloque 3 (M=0,57) (Bq1>Bq2>Bq3).

El test de normalidad Shapiro-Wilk ha revelado que los datos de las variables de la fase 3 de codificación no presentan una distribución normal (anexo 6, tabla 34 y tabla 35). Los resultados del test de Kruskal-Wallis indican que hay diferencia estadísticamente significativa entre los tres niveles respecto tanto a las variables de la fase 3a de codificación (sustantivo, adjetivo, verbo y adverbio) como en el caso de las variables de la fase 3b de codificación (errores sustantivo, errores adjetivo, errores verbo y errores adverbio) (valor-p<0,05) (anexo 7, tabla 42 y tabla 43).

Los datos obtenidos a partir de las pruebas estadísticas U de Mann-Whitney (tabla 15, tabla 16, tabla 17, tabla 18) revelan que no hay diferencia estadísticamente significativa respecto a las medias de errores en las clases de palabra léxica entre los bloques 1 y 2 (valor-p>0,05), excepto en el caso de la variable errores verbo (valor-p=0,035). Este resultado indica que la media de errores en los verbos disminuye de manera significativa entre los niveles inicial e intermedio. Sin embargo, la tendencia inicial que también se ha observado para las demás variables (Bq1>Bq2) no se confirma, es decir, la mejora no es significativa estadísticamente. Asimismo, cabe recordar que con respecto a la variable errores adjetivo, la tendencia inicial indicaba que los errores aumentaban entre los bloques 1 y 2 (Bq1<Bq2), pero como acabamos de ver, tampoco se verifica diferencia estadísticamente significativa en este caso.

Tabla 15. Test U de Mann-Whitney, variable errores sustantivo.

<b>Variable: errores sustantivo</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	123,21	-1,769	,077
2	108	107,79		
2	108	126,28	-,965	,335
3	134	117,65		
1	123	142,77	-2,874	<b>,004</b>
3	134	116,36		

Tabla 16. Test U de Mann-Whitney, variable errores adjetivo.

<b>Variable: errores adjetivo</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	115,66	-,084	,933
2	108	116,38		
2	108	131,39	-2,060	<b>,039</b>
3	134	113,53		
1	123	138,65	-2,082	<b>,037</b>
3	134	120,14		

Tabla 17. Test U de Mann-Whitney, variable errores verbo.

<b>Variable: errores verbo</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	124,62	-2,104	<b>,035</b>
2	108	106,19		
2	108	141,00	-3,930	<b>,000</b>
3	134	105,78		
1	123	159,72	-6,396	<b>,000</b>
3	134	100,80		

Tabla 18. Test U de Mann-Whitney, variable errores adverbio.

<b>Variable: errores adverbio</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	119,72	-,967	,333
2	108	111,76		
2	108	130,21	-1,942	<b>,05</b>
3	134	114,48		
1	123	142,72	-3,109	<b>,002</b>
3	134	116,40		

Los resultados de esta prueba para la diferencia observada entre los bloques de análisis 2 y 3 indican que en el caso de la variable errores sustantivo (valor-p=0,335) no se verifica significación estadística. En lo que se refiere a las demás variables (errores adjetivo, errores verbo y errores adverbio), sí hay diferencia estadísticamente significativa (valor-p<0,05), lo que confirma la tendencia observada de que los textos del nivel avanzado presentan menos errores que los del nivel intermedio (Bq2>Bq3).

Finalmente, los resultados de la prueba de Mann-Whitney revelan que hay diferencia estadísticamente significativa entre los bloques de análisis 1 y 3 para todas las variables (errores sustantivo, errores adjetivo, errores verbo y errores adverbio) (valor-p<0,05).

En resumen, la clase de palabra léxica más abundante es la de los sustantivos, seguida de la de los verbos. Sin embargo, la clase que presenta más errores es la del **verbo**, seguida por la de los **sustantivos**. En el caso de los sustantivos, sólo se puede decir que se ha verificado una disminución significativa de los errores entre los niveles inicial y avanzado. Esto puede ser una característica propia del aprendizaje de lenguas próximas que indica que son necesarias más horas de instrucción para que el alumno utilice los sustantivos correctamente. En cambio, en lo que se refiere a los errores identificados en los verbos, vemos que hay disminución significativa entre todos los niveles. Con respecto a los **adjetivos** y a los **adverbios**, se nota que el número de errores baja significativamente solamente entre los niveles intermedio y avanzado.

#### 10.4. Resultados y análisis del tipo de error léxico

**Pregunta 4: ¿Qué tipo de errores léxicos (formal, semántico o mixto) se observan en los textos escritos por profesores brasileños de ELE en formación según el número de horas de instrucción?**

En este apartado analizamos los resultados referentes a los datos identificados en la fase 4 de codificación, es decir, el tipo de error en las palabras léxicas. Como se ha visto en la tabla 5, los 3.302 errores identificados en este grupo de palabras se han clasificado en errores que afectaban a la forma de la palabra léxica (errores formales), al contenido de la palabra léxica (errores semánticos) o errores que afectaban tanto a la forma como al contenido de la palabra léxica (errores mixtos).

A simple vista se nota que el tipo de error más abundante es el error formal (N=2.427) lo que corresponde al 73,5% de los errores. Los errores semánticos (N=467) afectan al 14,14% del total de errores, seguido de los errores mixtos (N=408), correspondientes al 12,36% de los errores en las palabras léxicas. En la tabla 19 presentamos algunos ejemplos de estos errores que se comentan con más detenimiento en el capítulo dedicado a la discusión de los resultados.

Tabla 19. Ejemplos de errores léxicos según el tipo.

Errores léxicos según el tipo					
formales		semánticos		mixtos	
error (interlengua)	español (L2)	error (interlengua)	español (L2)	error (interlengua)	español (L2)
coerente*	coherente	Cabo se quedó*	pasó a ser/ fue nombrada	bailar envuelta* de una hoguera	alrededor de
perceber*	percibir	La región más animada*	divertida	criança*	niño

Con el objetivo de analizar los datos en profundidad y verificar si se observa un cambio entre los tres bloques de análisis respecto al tipo de error a medida que los alumnos avanzan en su período de formación, pasamos a analizar las medias de cada tipo de error por texto en cada uno de los niveles. La media de **errores formales** identificados en el bloque de análisis 1 (M=8,21) es superior a la media del bloque 2 (M=6,71) y que ésta es superior a la media del bloque 3 (M=5,16). Se observa, por tanto, una disminución de errores de tipo formal a medida que aumenta el número de horas de instrucción (Bq1>Bq2>Bq3) (figura 17).

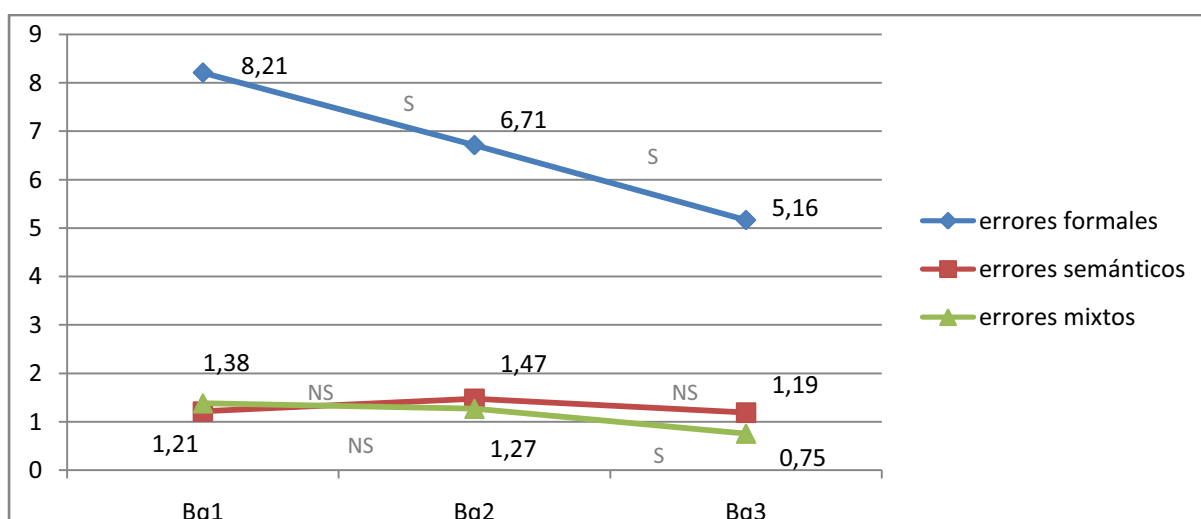


Figura 17. Media del tipo de error identificado por texto.

Nota: S=diferencia significativa; NS=diferencia no significativa.

Por lo que respecta a los **errores semánticos**, detectamos que esta variable no sigue la tendencia observada en los errores formales. En este caso, la media de errores semánticos del bloque 2 (M=1,47) es superior tanto a la media del bloque 1 (M=1,21) como a la del bloque 3 (M=1,19) (Bq1<Bq2>Bq3), aunque vemos que estas diferencias entre las medias

son bastante discretas. En cuanto a los **errores mixtos**, siguen la tendencia observada en los errores formales ( $Bq1 > Bq2 > Bq3$ ), es decir, la media de errores identificados en el bloque 1 ( $M=1,38$ ) es superior a la media encontrada en el bloque 2 ( $M=1,27$ ) y ésta es superior a la media encontrada en el bloque 3 ( $M=0,75$ ), bajando de manera progresiva pero discreta.

A continuación realizamos las pruebas estadísticas de comparación de medias para verificar si las diferencias son significativas. Como en las variables anteriores, procedemos a realizar el test de normalidad de Shapiro-Wilk, que reveló que los datos de las variables referentes al tipo del error tampoco presentan una distribución normal (valor- $p < 0,05$ ) (anexo 6, tabla 36).

Para verificar si hay diferencia entre los tres bloques de análisis, recurrimos a la prueba de Kruskal-Wallis (anexo 7, tabla 44), cuyos resultados indican que en el caso de los errores formales (valor- $p=0,000$ ) y mixtos (valor- $p=0,000$ ) existe una diferencia estadísticamente significativa entre los tres bloques, pero no en el caso de los errores semánticos (valor- $p=0,697$ ).

Como en los casos anteriores, aplicamos la prueba U de Mann-Whitney para conocer entre qué bloques de análisis se verifica esta diferencia. En la tabla 20 vemos los resultados del test U de Mann-Whitney que revelan que en el caso de los **errores formales**, la diferencia observada inicialmente tanto entre los bloques 1 y 2 (valor- $p=0,035$ ) como entre los bloques 2 y 3 (valor- $p=0,003$ ) es estadísticamente significativa. Como era de esperar, también se verifica diferencia estadísticamente significativa entre los bloques 1 y 3 (valor- $p=0,000$ ). Estos resultados nos permiten afirmar que la media de errores formales en las palabras léxicas disminuye a medida que aumenta el número de horas de instrucción.

Tabla 20. Test U de Mann-Whitney: variable errores formales.

Variable errores formales				
Bloques	N	Rango medio	Z	Valor-p
1	123	124,65	-2,104	<b>,035</b>
2	108	106,15		
2	108	136,50	-3,003	<b>,003</b>
3	134	109,41		
1	123	155,98	-5,593	<b>,000</b>
3	134	104,23		

Tal y como se ha observado anteriormente, no se verifica diferencia estadísticamente significativa para la variable **errores semánticos** entre los tres niveles (tabla 21), es decir, no

se observa una mejora de este tipo de error a lo largo del período de formación (valor- $p > 0,05$ ). Es importante profundizar en el análisis de estos errores para averiguar si se identifica una tendencia a la fosilización.

Tabla 21. Test U de Mann-Whitney: variable errores semánticos.

Variable errores semánticos				
Bloques	N	Rango medio	Z	Valor-p
1	123	115,20	-,203	,839
2	108	116,91		
2	108	125,34	-,799	,424
3	134	118,41		
1	123	131,89	-,624	,533
3	134	126,35		

En el caso de la variable **errores mixtos**, no se observa diferencia estadísticamente significativa para esta variable entre los bloques 1 y 2 (valor- $p = 0,423$ ), por lo tanto, se rechaza la hipótesis nula de que los bloques son diferentes. Sin embargo, entre los bloques 2 y 3 (valor- $p = 0,001$ ) y entre los bloques 1 y 3 (valor- $p = 0,000$ ) la diferencia es estadísticamente significativa (tabla 22).

Tabla 22. Test U de Mann-Whitney: variable errores mixtos.

Variable errores mixtos				
Bloques	N	Rango medio	Z	Valor-p
1	123	119,17	-,802	,423
2	108	112,39		
2	108	136,55	-3,225	,001
3	134	109,37		
1	123	147,52	-4,075	,000
3	134	112,00		

En resumen, los errores formales son mucho más numerosos que los errores semánticos o mixtos. Sin embargo, aunque los errores formales disminuyan significativamente a medida que aumenta el número de horas de instrucción, siguen siendo los más abundantes en los textos del nivel avanzado. Por otro lado, el hecho de que no se observe diferencia estadísticamente significativa entre los niveles con relación a los errores semánticos nos permite notar que aunque estos errores sean mucho menos frecuentes que los errores formales, son persistentes y están presentes incluso cuando los estudiantes ya han alcanzado un nivel de competencia avanzado.

## 10.5. Resultados y análisis del origen del error léxico

**Pregunta 5: ¿Cuál es el origen de los errores léxicos en la producción escrita de los estudiantes brasileños de ELE a lo largo del período de formación? ¿Los errores se deben a la influencia de su L1 (error interlingüístico), el portugués variedad brasileña, o son fruto del proceso de desarrollo de la interlengua (error intralingüístico)?**

Para contestar esta pregunta analizamos los datos referentes a la fase 5 de codificación (tabla 5), en la que se etiquetaron las palabras léxicas en función del origen o la causa del error. Los errores de origen interlingüístico (L1, portugués variedad brasileña) son los más frecuentes (N=1.644; 49,79%), seguido de los errores de origen intralingüístico (L2, español) (N=1.487; 45,03%). Los errores de origen ambiguo, es decir, aquellos errores cuyo origen no se pudo precisar, son muy poco frecuentes (N=171; 5,18%). En la tabla 23 presentamos algunos ejemplos de los errores identificados según el origen.

Tabla 23. Ejemplos de errores léxicos según el origen.

Errores léxicos según el origen								
interlingüístico			intralingüístico			ambiguo		
portugués (L1)	error (IL)	español (L2)	portugués (L1)	error (IL)	español (L2)	portugués (L1)	error (IL)	español (L2)
<i>compreendo</i>	compreendo*	comprendo	<i>passatempo</i>	aficción*	afición	<i>ensinos</i>	enseños*	enseñanzas
<i>prazer</i>	plazer*	placer	<i>tive</i>	tuvo*	tuve	<i>quisessem</i>	queraían*	quisieran
<i>projetos</i>	projectos*	proyectos	<i>exame, prova</i>	examen*	examen	<i>inesquecível</i>	inesquecible*	inolvidable

En la figura 18 presentamos las medias de errores por texto según el origen en cada uno de los bloques. En el caso de los errores interlingüísticos la media del bloque 1 (M=5,94) es mayor que la media del bloque 2 (M=4,13) y que ésta es mayor que la media del bloque 3 (M=3,49) (Bq1>Bq2>Bq3). En el caso de los errores intralingüísticos, la media de errores del bloque 1 (M=4,34) es menor que la media de errores en el bloque 2 (M=4,74) (Bq1<Bq2). Sin embargo, la media de errores del bloque 3 (M=3,29) es menor que la media del bloque 2 (Bq2>Bq3) y también que la del bloque 1 (Bq1>Bq3). Finalmente, vemos que en el caso de los errores ambiguos ocurre algo similar porque la media de errores del bloque 2 (M=0,58) es superior a la de los bloques 1 (M=0,52) y 3 (M=0,33) (Bq1<Bq2>Bq3).

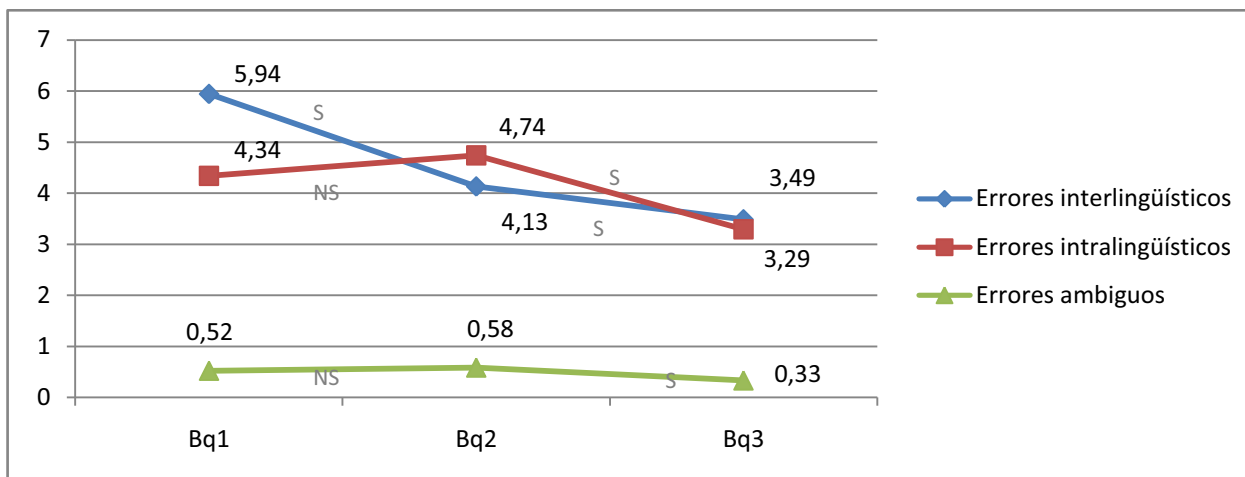


Figura 18. Media del número de errores según el origen en cada bloque de análisis.  
 Nota: S=diferencia significativa; NS=diferencia no significativa.

A continuación, tal y como se hizo con las variables anteriores, realizamos pruebas estadísticas para confirmar si las diferencias observadas son significativas. El test de normalidad Shapiro-Wilk confirmó que los datos de las variables correspondientes al origen del error léxico no presentan una distribución normal (valor- $p < 0,05$ ) (anexo 6, tabla 37). Los resultados del test de Kruskal-Wallis indican que hay diferencia estadísticamente significativa entre los tres bloques (valor- $p < 0,05$ ) (anexo 7, tabla 45).

En la tabla 24 presentamos los resultados del test U de Mann-Whitney. En lo que se refiere a los errores interlingüísticos, vemos que hay diferencia estadísticamente significativa entre los tres niveles (valores- $p < 0,05$ ). Sin embargo, aunque se observe una mejora de los errores influenciados por la L1, vemos que éstos siguen siendo los errores según el origen más abundantes en el nivel avanzado.



Tabla 24. Test U de Mann-Whitney: variable origen del error.

<b>Variable errores interlingüísticos</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	129,57	-3,310	<b>,001</b>
2	108	100,55		
2	108	131,78	-2,065	<b>,039</b>
3	134	113,22		
1	123	153,97	-5,186	<b>,000</b>
3	134	106,08		
<b>Variable errores intralingüísticos</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	114,22	-,435	,664
2	108	118,03		
2	108	138,79	-3,472	<b>,001</b>
3	134	107,56		
1	123	146,55	-3,654	<b>,000</b>
3	134	112,89		
<b>Variable errores ambiguos</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	115,44	-,156	,876
2	108	116,63		
2	108	132,01	-2,591	<b>,010</b>
3	134	113,03		
1	123	138,96	-2,539	<b>,011</b>
3	134	119,86		

En el caso de los errores intralingüísticos, vemos que el aumento de la media de estos errores que se ha observado anteriormente entre los bloques 1 y 2 ( $Bq1 < Bq2$ ) no es estadísticamente significativo (valor-p=0,664) (tabla 24). Sin embargo, se verifica diferencia estadísticamente significativa entre los bloques 2 y 3 (valor-p=0,001) y también entre los bloques 1 y 3 (valor-p=0,000). Esto revela que los estudiantes de los niveles avanzados producen menos errores de origen intralingüístico que los de los niveles intermedio.

En resumen, la mayor parte de los errores presentes en la muestra es de origen interlingüístico. Además, verificamos que la frecuencia tanto de los errores interlingüísticos como de los intralingüísticos disminuye de manera significativa a lo largo del proceso de aprendizaje. En el primer caso, la mejora se observa a lo largo de todo el proceso, mientras que en el segundo, observamos progreso solamente a partir del nivel avanzado. Los errores ambiguos son mucho menos frecuentes que los anteriores y también sólo disminuyen significativamente a partir del nivel avanzado.

## 10.6. Resultados y análisis según el tipo y el origen del error

**Pregunta 6: ¿Cambia el tipo del error léxico (formal, semántico o mixto) según el origen (interlingüístico, intralingüístico o ambiguo) en la producción escrita de los estudiantes brasileños durante el período de formación?**

Para contestar esta pregunta analizamos los datos referentes a las fases 4 y 5 de codificación (tabla 5). Dado que todos los errores están clasificados según el origen (interlingüístico, intralingüístico o ambiguo) y el tipo (formal, semántico o mixto), hemos considerado que un análisis de los mismos que tenga en cuenta la combinación de las dos variables permitiría profundizar en el estudio de los errores léxicos en un contexto de lenguas afines como es el caso del español (L2) y el portugués (L1).

Considerando el *corpus* general, llama la atención que los errores formales, tanto los de origen interlingüístico (35,3%) como intralingüístico (35,7%), son los que tienen un mayor peso en la muestra (figura 19). Por otro lado, los errores semánticos tanto interlingüísticos (7,6%) como intralingüísticos (6,2%) tienen una presencia bastante discreta en comparación con los formales. Asimismo, los errores mixtos tienen diferente peso según el origen del error, siendo más numerosos en el caso de los interlingüísticos (6,9%) que en el de los intralingüísticos (3,1%). Finalmente, se nota que los errores de origen ambiguo son muy poco frecuentes y que los de tipo semántico (0,3%) casi no se producen en comparación con los formales (2,5%) y mixtos (2,3%).

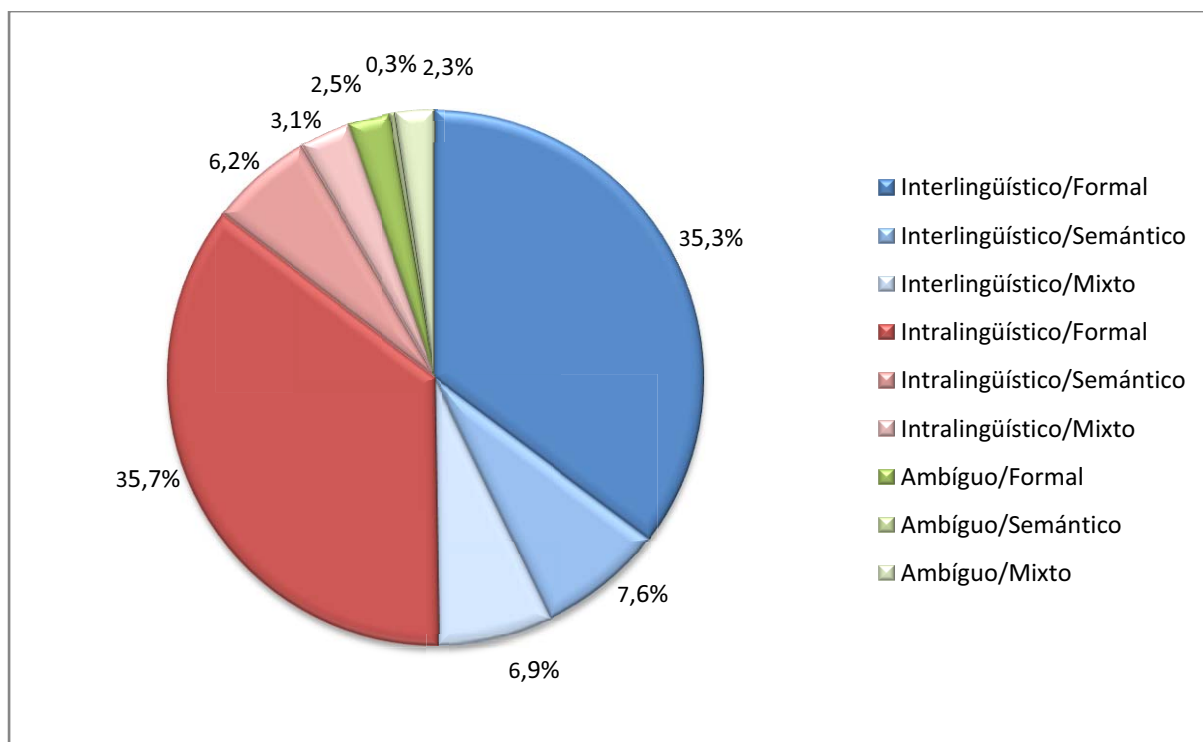


Figura 19. Errores léxicos según el tipo y el origen.

A continuación, presentamos algunos ejemplos de esos errores que se han identificado en el *corpus*: error interlingüístico según el tipo (tabla 25), error intralingüístico según el tipo (tabla 26) y error ambiguo según el tipo (tabla 27).

Tabla 25. Ejemplos de errores léxicos de origen interlingüístico según el tipo.

		Ejemplo	Contexto
Error interlingüístico	Formal	dificultad*: dificultad (esp.) <i>dificuldade</i> (port.)	Tal vez por eso tengo un poco de <b>dificultad*</b> con los vocábolos* (...)
		vocábolos*: vocablos (esp.) <i>vocábulos</i> (port.)	<b>Quando*</b> enpecé* a estudiar español (...)
		quando*: cuando (esp.) <i>quando</i> (port.)	
	Semántico	pesquisas*: investigaciones (esp.) <i>pesquisas</i> (port.)	(...) penso* en la enseñanza de la lengua y en trabajar con <b>pesquisas*</b> y estudios.
		serie*: curso (esp.) <i>série</i> (port.)	Yo estaba haciendo la última <b>serie*</b> de la enseñanza fundamental.
		investir*: invertir (esp.) <i>investir</i> (port.)	(...) pero no tienen dinero para <b>investir*</b> en este campo.
	Mixto	cumprimientos*: felicitaciones (esp.) <i>cumprimentos</i> (port.)	Al presidente de Brasil (...), mis <b>cumprimientos*</b> .
		vacas*: plazas (esp.) <i>vagas</i> (port.)	(...) el aumento* de <b>vacas*</b> en los hospitales.
		cualidad*: calidad (esp.) <i>qualidade</i> (port.)	(...) educación básica de <b>cualidad*</b> .

\* Se refieren a los ejemplos de palabras léxicas con error que se han identificado y seleccionado con el fin de ilustrar la clasificación utilizada y enriquecer la información obtenida en el análisis cuantitativo.

Tabla 26. Ejemplos de errores léxicos de origen intralingüístico según el tipo.

		Ejemplo	Contexto
Error intralingüístico	Formal	países*: países (esp.) <i>países</i> (port.) aficción*: afición (esp.) <i>passatempo</i> (port.) piensado*: pensado (esp.) <i>pensado</i> (port.)	(...) países* latino americanos*. (...) gente desconocida pero unidas por la misma aficción*. Yo nunca havia pensado* en estudiar español (...)
	Semántico	peleaban*: reñían (esp.) <i>reprendiam</i> (port.) arreglarse*: conseguir (esp.) <i>arranjar</i> (port.) acordar*: recordar (esp.) <i>lembrar</i> (port.)	(...) las profesoras (...) nos peleaban* cuando necesario. (...) con la finalidad de arreglarse* nobios*. (...) pero tenemos que acordar* que el voto libre y la democracia (...)
	Mixto	prójima*: próxima (esp.) <i>próxima</i> (port.) juciosas*: jugosas (esp.) <i>suculentas</i> (port.) me hecho* de menos: echo de menos (esp.) <i>sinto falta de</i> (port.)	(...) la prójima* vez. (...) frutas juciosas* e* deliciosas. Me hecho* de menos este lugar.

\* Se refieren a los ejemplos de palabras léxicas con error que se han identificado y seleccionado con el fin de ilustrar la clasificación utilizada y enriquecer la información obtenida en el análisis cuantitativo.

Tabla 27. Ejemplos de errores léxicos de origen ambiguo según el tipo.

		Ejemplos	Contexto
Error ambiguo	Formal	aprofundado*: ahondado/a fondo (esp.) <i>aprofundado</i> , profundo (port.) vergonza*: vergüenza (esp.) <i>vergonha</i> (port.) diñero*: dinero (esp.) <i>dinheiro</i> (port.)	(...) puedo hacer un estudio más aprofundado* de la lengua. (...) me quede* rojo de vergonza*. (...) hacer cursos de español solamente por diñero* (...)
	Semántico	se quedase*: se volviera (esp.) <i>ficase</i> (port.) se tornó*: se convirtió (esp.) <i>se tornou</i> (port.) atrapa*: impide/dificulta (esp.) <i>impede/dificulta</i> (port.)	Hicieron con que el clima se quedase* difícil*. (...) mi mejor amiga se tornó* mi madriña. (...) la ley atrapa* el trabajo del profesor.
	Mixto	afuera*: aparte de (esp.) <i>fora/além</i> de (port.) echa*: creo (esp.) acho (port.) flaco*: débil (esp.) <i>fraco</i> (port.)	(...) afuera* eso a mí me gusta porque (...) (...) gusto mucho del español y echa* una lengua muy guapa. (...) sus* contenidos ministrado* eran muy flaco*.

\* Se refieren a los ejemplos de palabras léxicas con error que se han identificado y seleccionado con el fin de ilustrar la clasificación utilizada y enriquecer la información obtenida en el análisis cuantitativo.

Pasamos a presentar y analizar las medias de esos errores por texto en cada uno de los niveles. Cabe destacar que los errores formales tanto interlingüísticos como intralingüísticos

siguen la tendencia anteriormente observada y son los más abundantes en cada uno de los tres niveles, aunque con distinta proporción en cada uno de ellos (figura 20).

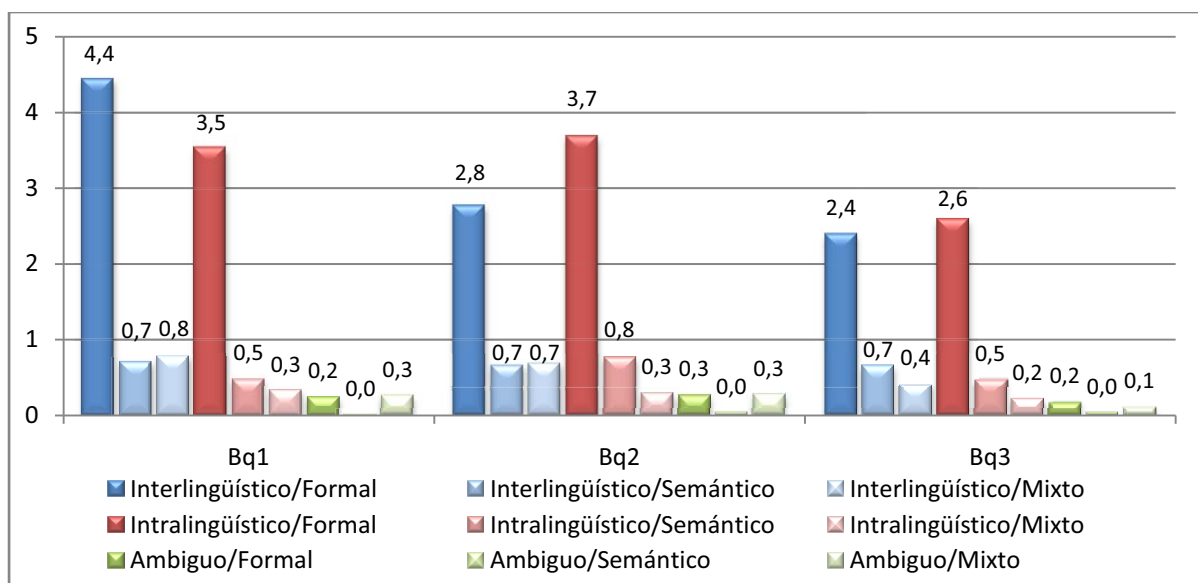


Figura 20. Media de errores léxicos por texto según el origen y el tipo.

Tras presentar una visión general del comportamiento de esos errores, pasamos a averiguar si hay cambios destacables a lo largo del período de formación tal como hemos hecho hasta el momento.

En primer lugar analizamos los errores de origen interlingüístico según el tipo (figura 21). Tal como se ha visto más arriba en el *corpus* general, los errores formales son los más abundantes a lo largo de los tres niveles. Con todo, el análisis de las medias por bloques revela que éstas disminuyen a medida que aumenta el número de horas de instrucción (Bq1>Bq2>Bq3). Con relación a los errores semánticos, observamos que la media del bloque 1 (M=0,72) es ligeramente superior a la de los bloques 2 y 3 (M=0,67) (Bq1>Bq2=Bq3). Es interesante notar que las medias de los bloques 2 y 3 son idénticas. En el caso de los errores mixtos, es decir, aquellos que presentan incorrecciones tanto en la forma como en el significado, podemos decir que disminuyen a lo largo del proceso de aprendizaje (Bq1>Bq2>Bq3).

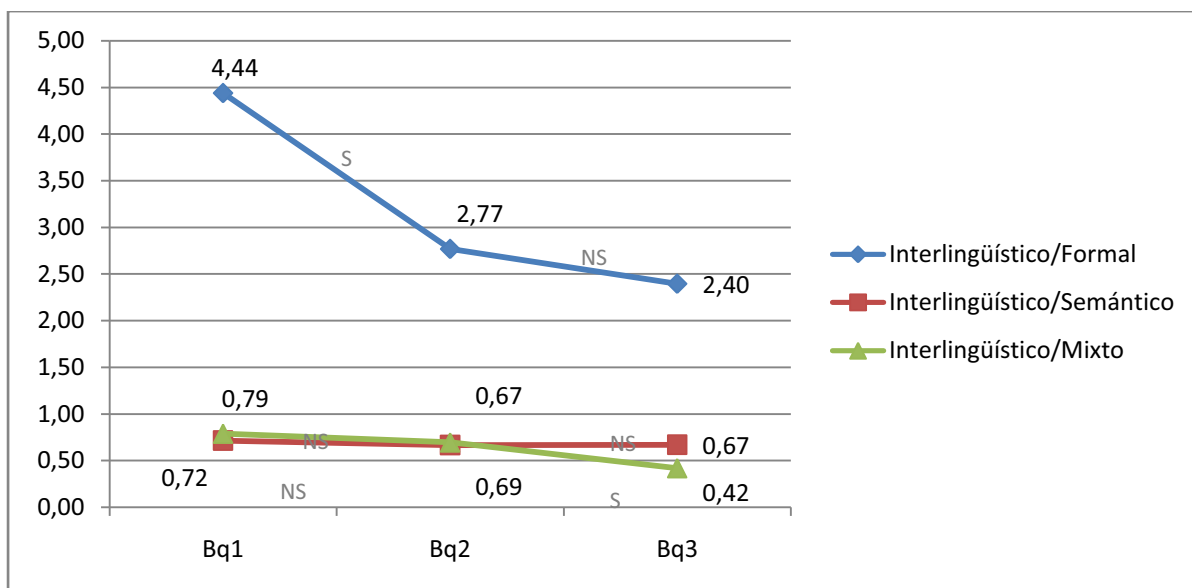


Figura 21. Media de errores interlingüísticos según el tipo.  
 Nota: S=diferencia significativa; NS=diferencia no significativa.

Tras la aplicación de las pruebas estadísticas de Mann-Whitney (tabla 28), comprobamos que la disminución que se observa a lo largo de los niveles en los errores formales de origen interlingüístico es estadísticamente significativa (valor- $p=0,000$ ) entre los bloques 1 y 2. Sin embargo, la disminución en la media de errores que se observa entre los bloques 2 y 3 no lo es (valor- $p=0,058$ ). En cuanto a los errores semánticos de origen interlingüístico, no se verifica diferencia estadísticamente significativa entre ninguno de los niveles (valor- $p>0,05$ ), lo que puede indicar una tendencia a la fosilización de este tipo de error. Por otro lado, en el caso de los errores mixtos también de origen interlingüístico, sí hay diferencia estadísticamente significativa entre los bloques 2 y 3 (valor- $p=0,017$ ).

Tabla 28. Test U de Mann-Whitney: variable error interlingüístico según el tipo.

<b>Variable: Error Interlingüístico/Formal</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	118,80	-3,544	<b>0,000</b>
2	108	112,81		
2	108	130,87	-1,895	0,058
3	134	113,95		
1	123	154,47	-5,314	<b>0,000</b>
3	134	105,62		
<b>Variable: Error Interlingüístico/Semántico</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	118,80	-0,766	0,444
2	108	112,81		
2	108	119,43	-0,468	0,640
3	134	123,17		
1	123	130,46	-0,336	0,737
3	134	127,66		
<b>Variable: Error Interlingüístico/Mixto</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	118,54	-0,688	0,491
2	108	113,11		
2	108	131,55	-2,396	<b>0,017</b>
3	134	113,40		
1	123	142,08	-3,168	<b>0,002</b>
3	134	116,99		

En segundo lugar, analizamos los errores de origen intralingüístico según el tipo (figura 22). Nuevamente destacamos el peso que tienen los errores formales frente a los errores semánticos y mixtos. Además, vemos que la media de errores formales del bloque 2 ( $M=3,68$ ), es superior a la de los bloques 1 ( $M=3,53$ ) y 3 ( $M=2,59$ ) ( $Bq1 < Bq2 > Bq3$ ). En el caso de los errores semánticos, las medias de los bloques 1 y 3 ( $M=0,48$ ) son inferiores a la del bloque 2 ( $M=0,77$ ) ( $Bq1 < Bq2 > Bq3$ ). Es interesante notar que la media del bloque 3 es idéntica a la media del bloque 1 ( $Bq1 = Bq3$ ). En cuanto a los errores mixtos, los menos numerosos, vemos que las medias de estos errores disminuyen a lo largo de los tres niveles ( $Bq1 > Bq2 > Bq3$ ).

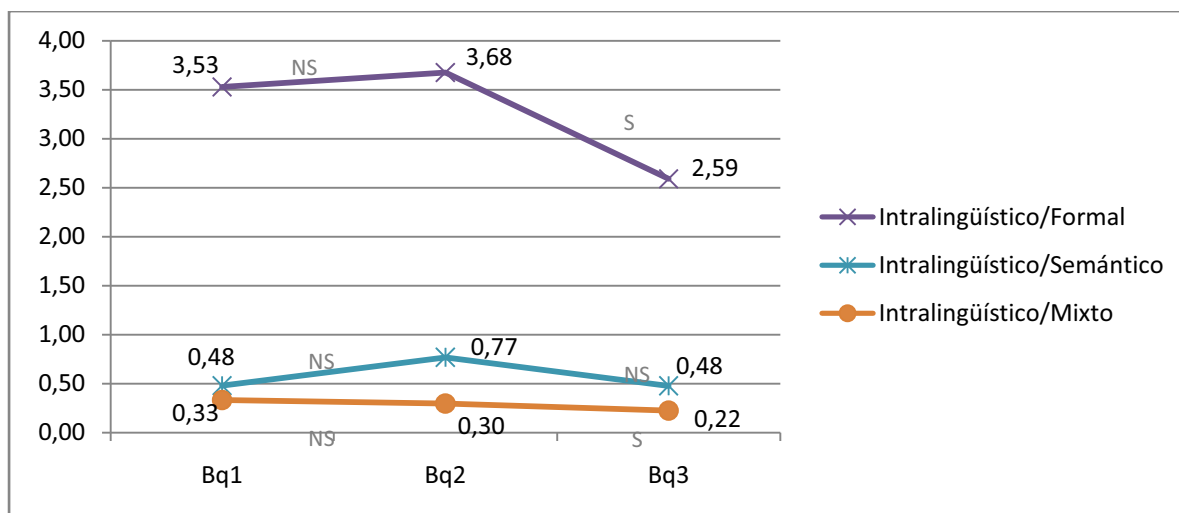


Figura 22. Media de los errores intralingüísticos según el tipo por bloque de análisis.

Nota: S=diferencia significativa; NS=diferencia no significativa.

Las pruebas estadísticas de Mann-Whitney (tabla 29) revelaron que en lo que se refiere a los errores formales de origen intralingüístico la disminución es estadísticamente significativa entre los bloques 2 y 3 (valor- $p=0,004$ ). Con respecto a los errores semánticos, no se ha observado diferencia estadísticamente significativa entre las medias de ninguno de los tres bloques, lo que puede indicar una vez más una tendencia a la fosilización de estos errores tal y como se ha observado más arriba en el análisis de los errores de origen interlingüístico. En cuanto a los errores mixtos de origen intralingüístico, solamente se ha observado diferencia estadísticamente significativa entre los bloques 2 y 3 (valor- $p=0,014$ ).



Tabla 29. Test U de Mann-Whitney: variable tipo del error intralingüístico.

<b>Variable: Error Intralingüístico/Formal</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	116,74	-0,180	0,857
2	108	115,16		
2	108	135,81	-2,882	<b>0,004</b>
3	134	109,97		
1	123	145,33	-3,407	<b>0,001</b>
3	134	114,01		
<b>Variable: Error Intralingüístico/Semántico</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	110,58	-1,497	0,134
2	108	122,18		
2	108	128,75	-1,653	0,098
3	134	115,66		
1	123	129,63	-0,153	0,878
3	134	128,42		
<b>Variable: Error Intralingüístico/Mixto</b>				
Bloques	N	Rango medio	Z	Valor-p
1	123	116,45	-0,142	0,887
2	108	115,49		
2	108	126,74	-1,474	<b>0,014</b>
3	134	117,28		
1	123	134,67	-1,640	0,101
3	134	123,80		

Finalmente, analizamos las medias de los errores ambiguos (figura 23). Con relación a los errores formales, el bloque de análisis 1 ( $M=0,24$ ) presenta una media inferior a la del bloque 2 ( $M=0,27$ ). Sin embargo, ésta es superior a la del bloque 3 ( $M=0,18$ ) ( $Bq1 < Bq2 > Bq3$ ). Al analizar las medias de los errores ambiguos que afectan a la forma, vemos que presentan un comportamiento similar al de los errores formales de origen intralingüístico, porque en ambos casos hay un aumento de la presencia de este error en el bloque 2 con relación al bloque 1. Con todo, se observa que las medias de errores del bloque 3 son inferiores incluso a las del bloque 1. Este comportamiento indica que, aunque haya un aumento en el número de errores en los niveles intermedio de aprendizaje, este tipo de error disminuye cuando los estudiantes alcanzan un nivel avanzado. Más adelante se verifica si esta disminución es estadísticamente significativa. Por otro lado, el análisis de las medias de los errores semánticos revela que en este caso son los menos frecuentes y que las medias de los bloques 2 y 3 son idénticas ( $M=0,04$ ) y superiores a la del bloque 1 ( $M=0,02$ ) ( $Bq1 < Bq2 = Bq3$ ). Este comportamiento de los errores semánticos de origen

ambiguo es el mismo que se ha visto anteriormente con los errores de origen interlingüístico, cuyas medias son idénticas en los bloques 2 y 3. Además, en el análisis de las medias de los errores semánticos de origen intralingüístico, hemos visto que la media del bloque 1 era igual a la del bloque 3. En lo que se refiere a los errores mixtos, éstos son más numerosos que los demás en los bloques 1 (M=0,26) y 2 (M=0,28), aunque en el bloque 3 (M=0,11) se observe una disminución acentuada de los mismos.

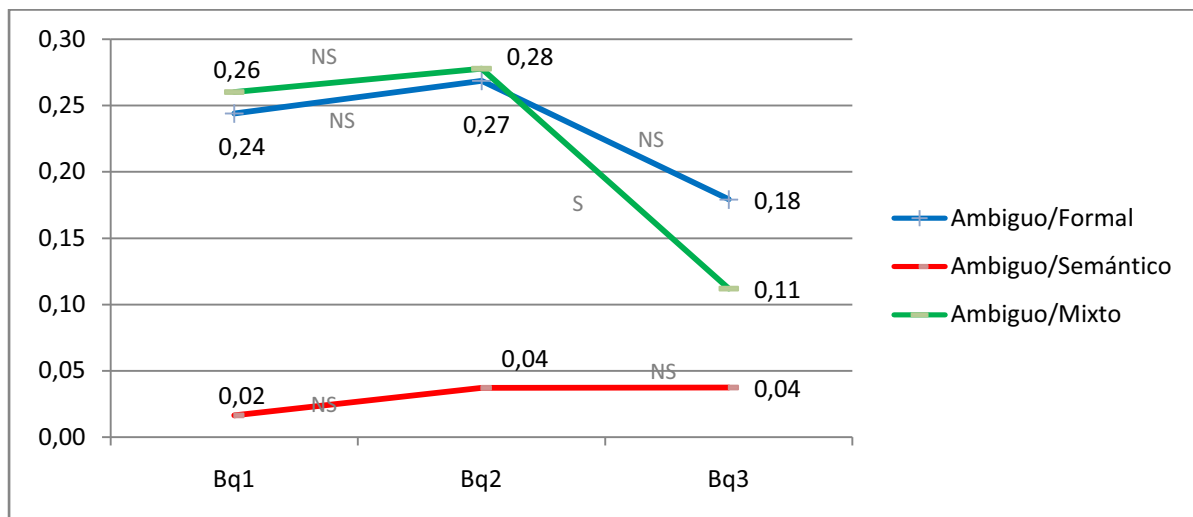


Figura 23. Media de los errores ambiguos según el tipo por bloque de análisis.

Nota: S=diferencia significativa; NS=diferencia no significativa.

En la tabla 30 vemos los resultados de la prueba estadística de Mann-Whitney que revela que solamente en el caso de los errores ambiguos mixtos, la disminución observada entre los bloques 2 y 3 es estadísticamente significativa (valor-p=0,017).

Tabla 30. Test U de Mann-Whitney: variable tipo del error ambiguo.

Variable: Error Ambiguo/Formal				
Bloques	N	Rango medio	Z	Valor-p
1	123	115,41	-0,201	0,840
2	108	116,67		
2	108	126,85	-1,639	0,101
3	134	117,19		
1	123	133,66	-1,485	0,138
3	134	124,72		
Variable: Error Ambiguo/Semántico				
Bloques	N	Rango medio	Z	Valor-p
1	123	114,88	-0,988	0,323
2	108	117,28		
2	108	121,48	-0,011	0,991
3	134	121,51		
1	123	127,59	-1,034	0,301
3	134	130,29		
Variable: Error Ambiguo/Mixto				
Bloques	N	Rango medio	Z	Valor-p
1	123	117,23	-0,409	0,683
2	108	114,60		
2	108	129,00	-2,394	<b>0,017</b>
3	134	115,46		
1	123	138,33	-2,952	<b>0,003</b>
3	134	120,43		

En resumen, tras el análisis de los errores léxicos según el origen y el tipo, podemos afirmar que aunque se observe una disminución significativa de los errores **interlingüísticos formales** a partir del nivel intermedio, esta disminución no es estadísticamente significativa en el nivel avanzado, lo que puede indicar una tendencia a la fosilización de este tipo de error. En el caso de los errores **interlingüísticos semánticos**, no se observa diferencia entre ninguno de los tres niveles. Como se ha dicho anteriormente, la presencia del error de tipo semántico es constante a lo largo del proceso de aprendizaje, aunque es menor que la de los errores formales. En lo que se refiere a los errores **interlingüísticos mixtos**, sólo se observa mejora a partir del nivel avanzado, lo que revela que los estudiantes necesitan más información sobre la L2 para superar este tipo de error.

El análisis de los errores **intra lingüísticos formales** indica que hay una mejora a partir del nivel avanzado. Una vez más vemos que el nivel de conocimiento sobre la lengua meta influye en la superación de algunos tipos de error en un contexto de lenguas próximas. En

cuanto a los errores **intralingüísticos semánticos**, una vez más no se observa ningún cambio entre los tres niveles analizados, lo que viene a reforzar la idea de que este tipo de error tiene una tendencia a fosilizarse. En el caso de los errores **intralingüísticos mixtos**, la mejora que se observa a partir del nivel avanzado no es significativa con respecto al nivel inicial.

Finalmente, el análisis de los errores de origen ambiguo sólo revela mejora en el caso de los errores **ambiguos mixtos** a partir del nivel avanzado. Además, se constata que este error es muy escaso y tanto en el caso de los errores **ambiguos formales** como **semánticos**, no se observa diferencia estadísticamente significativa entre los niveles.

### 10.7. Resultados y análisis de la percepción de la proximidad lingüística

Dado que la presente investigación se centra en conocer los errores léxicos a lo largo del período de formación de profesores de español en un contexto de lenguas afines (L2 español y L1 portugués), nos ha parecido pertinente explorar la percepción que tienen los participantes de la proximidad tipológica y el grado de facilidad o dificultad que les supone dicha proximidad cuando tienen que redactar en la lengua meta. Nos interesaba saber si los alumnos perciben la proximidad tipológica como un elemento facilitador y si la percepción de la facilidad o dificultad cambia según el nivel de competencia. Para ello, les solicitamos que contestaran algunas preguntas (anexo 4) al terminar de redactar los textos propuestos. Los resultados obtenidos se presentan a continuación.

Primeramente, destacamos que para el 56% (N=79) de los participantes les parece fácil escribir en español. Con todo, es interesante notar que esta percepción varía según el nivel de aprendizaje. En la figura 24 vemos que la percepción del ejercicio de la escritura como una actividad fácil aumenta a medida que los estudiantes avanzan en su aprendizaje, y aunque los datos indiquen que en el nivel inicial escribir en español es fácil solamente para el 44% los participantes, en los niveles intermedio (55%) y avanzado (64%) se observa un aumento en el número de estudiantes que comparten esta apreciación.

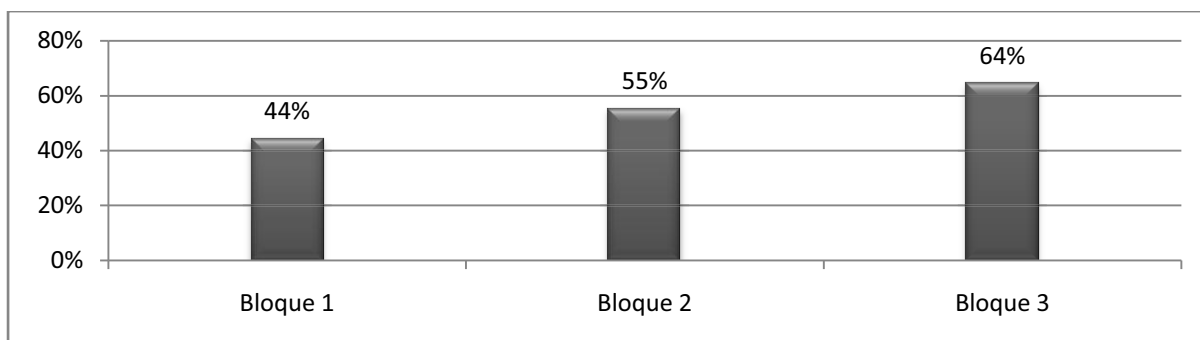


Figura 24. Percepción de los participantes sobre la facilidad de escribir en español.

Para justificar la percepción de la facilidad que sienten para escribir en español, los estudiantes citan la semejanza entre las dos lenguas (49%) y el hecho de que les gusta la lengua española (35%), sea por su sonoridad o por los referentes musicales y culturales. El 15% restante menciona otros factores como el conocimiento previo que tienen de la lengua meta, por ejemplo.

Paradójicamente, del 44% de participantes a los que les parece difícil escribir en español, el 75% destaca también la semejanza entre las dos lenguas como la razón principal de ello. Así, vemos que la proximidad tipológica se percibe por la mayoría de los participantes del estudio como un factor facilitador, pero se considera igualmente un elemento negativo que les produce dificultad a muchos de los alumnos cuando tienen que redactar en español.

En segundo lugar, quisimos conocer su percepción sobre la dificultad para redactar los textos propuestos (anexo 3). En este caso, el 61% (N=87) de los participantes afirmó haber tenido algún tipo de dificultad para producir las redacciones y señaló principalmente el vocabulario (50%) y la ortografía (21%) como razones de ello (figura 25).

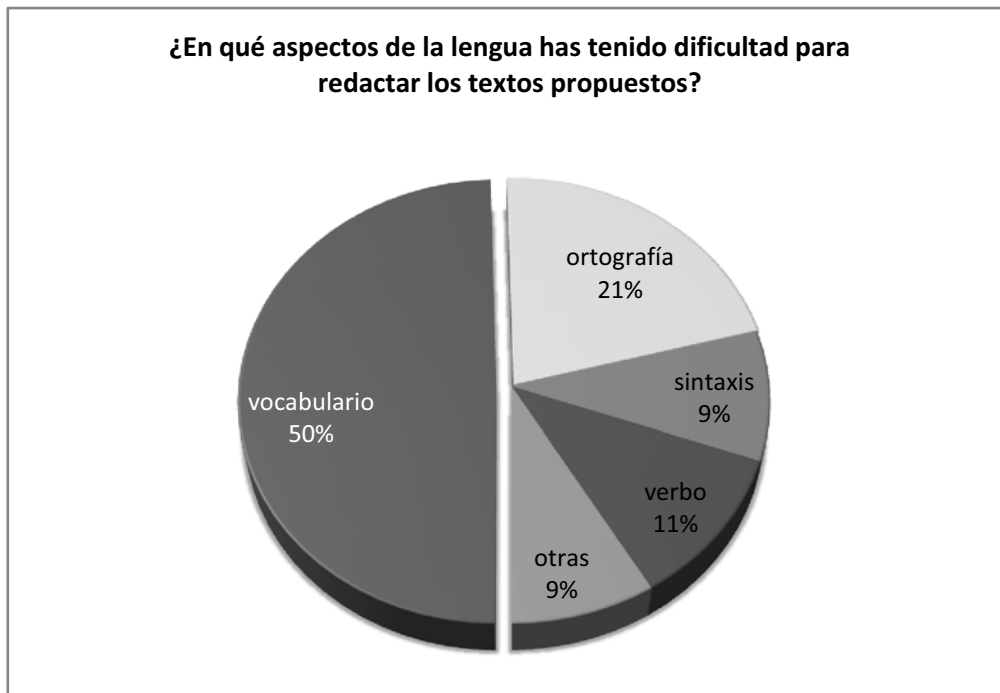


Figura 25. Razones de la dificultad para redactar los textos propuestos.

Los resultados del análisis cuantitativo corroboran los datos obtenidos a través del cuestionario de percepción de la proximidad lingüística porque, por un lado, los errores léxicos son los más numerosos (75%) frente a los funcionales (25%), lo que coincide con el hecho de que el vocabulario es el aspecto de la lengua más señalado por los participantes como factor de dificultad (50%); y por otro, el hecho de que los errores formales son los más abundantes (73,5%) y que este aspecto es el segundo más señalado como factor que les provoca dificultad (21%).

En tercer lugar, los participantes también autoevaluaron sus competencias en cada una de las destrezas o habilidades lingüísticas. En la figura 24 observamos que la mayoría de los participantes considera tener una competencia muy buena en las habilidades receptivas (el 47% en la comprensión oral y el 51% en la comprensión escrita) y en lo que corresponde a las habilidades productivas, la mayoría de los participantes considera tener tan sólo una competencia buena (el 36% en la expresión oral y el 37% en la expresión escrita).

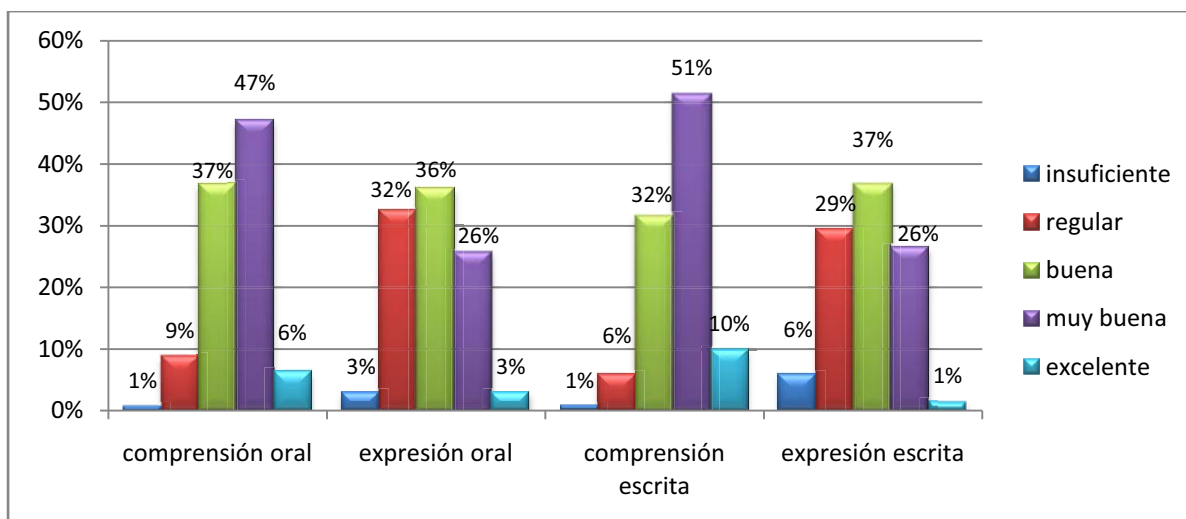


Figura 26. Autoevaluación de los participantes sobre las habilidades receptoras y productivas.

Finalmente, analizamos la percepción que tienen los participantes de sus capacidades en la expresión escrita, objeto de la presente investigación. En la figura 27 vemos que la percepción de los estudiantes cambia a medida que avanzan en el proceso de aprendizaje del español. Por un lado, se nota que en el nivel inicial (bloque 1), el 14% de los estudiantes considera tener un dominio insuficiente de esta destreza, en el nivel intermedio (bloque 2), este número baja a un 5%, y en el nivel avanzado (bloque 3) ninguno de los participantes ha señalado esta opción. Por otro, en el nivel inicial, la mayoría de los alumnos (45%) evalúa sus capacidades en esta destreza como regular, pero en el nivel intermedio (bloque 2), en cambio, vemos que la mayoría (38%) considera que tiene un buen dominio de la escritura, y finalmente, en el nivel avanzado (bloque 3), observamos el mismo resultado entre aquellos estudiantes que consideran que tienen tanto un buen (40%) como un muy buen dominio (40%) de la expresión escrita. Asimismo, los estudiantes demuestran ser muy críticos con respecto a sus habilidades en esta destreza ya que casi ninguno (2%) considera tener un nivel excelente a lo largo de los tres niveles de aprendizaje.

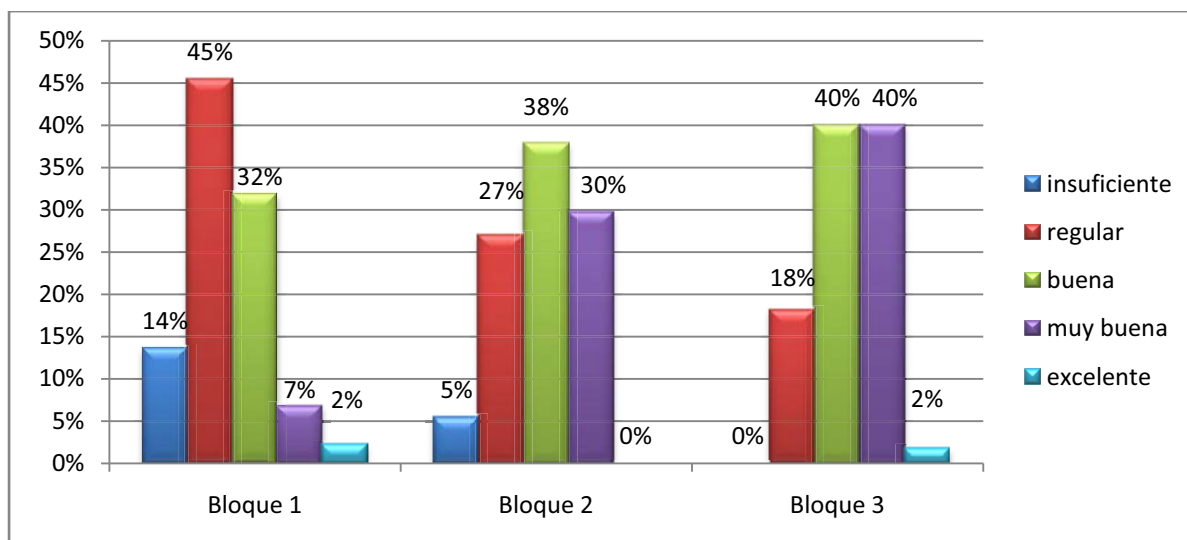


Figura 27. Autoevaluación de los participantes sobre la expresión escrita según el nivel.

En resumen, la mayoría de los alumnos percibe la proximidad lingüística como un elemento facilitador a partir del nivel intermedio y esta percepción aumenta en el nivel avanzado. Sin embargo, tanto los estudiantes que piensan que aprender español es fácil como los que piensan que es difícil citan la semejanza entre las lenguas como el principal motivo de su apreciación. Además, los estudiantes perciben que las destrezas expresivas les suponen más dificultad que las receptivas. Asimismo, en el caso específico de la expresión escrita, los estudiantes señalan que sus capacidades en esta destreza mejoran a medida que ellos avanzan en el proceso de aprendizaje de la L2.



## 11. DISCUSIÓN

En este capítulo se discuten los resultados obtenidos a partir del análisis de los datos. Para ello y para facilitar la lectura, hemos ordenado la discusión en diferentes apartados que se corresponden con las preguntas de investigación. Cada apartado está encabezado por una pregunta y un breve resumen de los resultados obtenidos que se han presentado en el capítulo anterior. A lo largo de la discusión ofrecemos ejemplos de los errores más frecuentes encontrados en la muestra. Recordamos que no es el objetivo de la presente investigación ofrecer listas de errores<sup>46</sup>, sino más bien, aplicando una metodología cuantitativa, presentar una visión general de las características de los errores léxicos según el tipo y el origen, y verificar si se observan cambios en cada una de las variables a lo largo del proceso de aprendizaje de español por estudiantes brasileños.

### 11.1. La fluidez y la diversidad léxica

---

**Pregunta 1** En la producción escrita de los profesores brasileños de ELE en formación ¿cambia la fluidez y la diversidad léxica a lo largo del período de formación?

**Resultados:** Sí. Hay un aumento relevante de la fluidez, es decir, los participantes son capaces de escribir textos más largos después de alcanzar el nivel intermedio, mientras que la mejora de la diversidad léxica sólo se verifica después que los alumnos alcanzan el nivel avanzado.

La fluidez y la diversidad léxica son medidas cuantitativas que se utilizan comúnmente para medir la calidad de lo escrito en diversos ámbitos de la investigación en Adquisición de Segundas Lenguas (ASL) (ej.: Wolfe-Quintero, *et al.*, 1998; Celaya *et al.*, 2001; Lasagabaster, 2001; Ortega, 2004; Torras *et al.*, 2006; Agustín Llach, 2011a). Cabe destacar que el uso de medidas cuantitativas como son la fluidez y la diversidad léxica para medir la calidad de los textos escritos es un recurso habitual en el ámbito del aprendizaje del inglés (L2) y que en el caso del aprendizaje de español como lengua extranjera (ELE) apenas existen.

Nuestro interés en analizar tanto la fluidez como la diversidad léxica reside en conocer el comportamiento de la interlengua de los estudiantes brasileños de ELE en tres momentos diferentes del período de formación para saber si en un contexto de lenguas próximas se

---

<sup>46</sup> Para listas de errores más detalladas se recomienda la consulta de Torijano Pérez (2002) y Durão (2004).

produce, por un lado, aumento en el número de palabras utilizadas en las redacciones y, por otro, si se observa un aumento en la variedad de palabras que son capaces de utilizar tras alcanzar diferentes niveles de competencia. Creemos que este conocimiento puede ayudar tanto a los estudiantes como a los profesores a entender el proceso de adquisición en estos contextos específicos y saber cuándo y en qué aspectos es necesario focalizar para que el aprendizaje sea más eficaz.

La fluidez se ha medido a través del número de palabras total presente en cada uno de los textos escritos (N total de palabras) y la diversidad léxica a través de la medida Uber U sin errores. Tras el análisis de los resultados, se ha visto que tanto la fluidez como la diversidad léxica cambian a lo largo del período de formación, es decir, aumentan a medida que aumenta el número de horas de instrucción. La fluidez aumenta significativamente entre los niveles inicial e intermedio, pero no se percibe un aumento significativo de la fluidez entre los niveles intermedio y avanzado, al contrario de lo que se ha observado en investigaciones anteriores en las que las lenguas analizadas no eran tipológicamente similares (Fernández, 1997; Palapanidi, 2010). El hecho de que la fluidez no aumente significativamente en el nivel avanzado con respecto al intermedio en la presente investigación puede deberse a que los estudiantes, dada la proximidad lingüística, consideren que han adquirido ya un nivel satisfactorio de la L2 que les permite comunicarse en la lengua meta y no sientan la necesidad de profundizar en el estudio de la lengua. En esta línea, Marangon (2011) apunta que la proximidad lingüística entre la L1 del estudiante y la L2 influye en el proceso de aprendizaje y ejemplifica, a través de la comparación del proceso de aprendizaje de italiano por estudiantes polacos y españoles, que los primeros tendrán claramente más dificultad al principio, pero en un nivel de competencia alto se observará mayor fluidez de vocabulario en su producción que en la de los estudiantes hispanohablantes. Asimismo, otro factor que creemos que ha influido en nuestros resultados se debe a las consignas que se dieron a los participantes. Estas consignas indicaban el número de palabras que deberían escribir en cada texto y tenían el objetivo de estimular la producción de las redacciones. Con todo, aunque los estudiantes del nivel inicial no siguieron las indicaciones, ya que escribieron menos de lo que se les había recomendado, constatamos que los participantes de los niveles intermedio y avanzado sí cumplieron con las consignas, lo que puede haber influido y limitado la fluidez en su producción escrita. Es por ello que sería necesario desarrollar un

estudio en el que no se limitara la producción de los estudiantes, y así confirmar si realmente ocurre un estancamiento de la fluidez tras el nivel intermedio en un contexto de lenguas afines.

En el caso de la diversidad léxica, los resultados obtenidos a través de la medida Uber U (Dugast, 1978; Jarvis, 2002) revelaron que ésta era más alta en los niveles iniciales. Con todo, el análisis más detallado de los errores encontrados en los textos escritos reveló que estos resultados se debían a la gran variación en la representación ortográfica de una misma palabra que se observó especialmente al principio del período de formación, es decir, en el nivel inicial. Con el objetivo de solucionar este problema, volvimos a analizar la diversidad léxica a través de la medida Uber U, pero esta vez excluyendo los errores previamente identificados (Engber, 1995). Los resultados de esta prueba indicaron que se da un aumento significativo de la diversidad léxica a lo largo del período de formación solamente cuando los participantes alcanzan el nivel avanzado. De ello se puede inferir que, en un contexto de lenguas próximas, es necesario un mayor número de horas de instrucción y contacto con la lengua meta para que la producción escrita refleje un uso más diversificado del vocabulario. De todos modos, los resultados de esta prueba quedan afectados tanto por la inclusión como por la exclusión de los errores identificados en la muestra porque, como se verá en el apartado siguiente, la presencia de los errores disminuye en el nivel avanzado.

En un estudio reciente sobre la utilización de medidas cuantitativas como índice de la calidad de lo escrito, Mellor (2011) encontró una gran correlación entre los resultados de la evaluación de la calidad de textos escritos en inglés (L2) con el análisis de la fluidez y la diversidad léxica. Sin embargo, en nuestro estudio, aunque la fluidez y la diversidad aumenten a medida que los estudiantes progresan en su período de formación, no observamos paralelismo entre estas medidas, ya que la fluidez mejora en el nivel intermedio y la diversidad sólo en el nivel avanzado. Aunque la fluidez no aumente en el nivel avanzado, ésta se ve compensada por un aumento de la diversidad léxica, es decir, los alumnos del nivel avanzado no escriben más que los de nivel intermedio, pero utilizan un vocabulario más variado.

A continuación se discuten los resultados del análisis de errores según la categoría gramatical y se comenta cómo se relacionan con la fluidez y la diversidad léxica.

## 11.2. Los errores léxicos y funcionales

---

**Pregunta 2** ¿En qué categoría de palabra (léxica o funcional) se observa un mayor número de errores según el número de horas de instrucción?

**Resultados:** El número de errores léxicos es superior al número de errores funcionales. El número de errores léxicos y funcionales presentes por texto disminuye según aumenta el número de horas de instrucción. Sin embargo, la disminución sólo es estadísticamente significativa entre los niveles intermedio y avanzado.

El interés por el análisis de los errores presentes en la interlengua de los estudiantes de segundas lenguas ha vuelto a cobrar importancia en los últimos años. En el campo del estudio de la adquisición del español como lengua extranjera (ELE), tras las valiosas aportaciones de las investigaciones de Vázquez (1991), Santos Gargallo (1993) y Fernández (1997), podemos nombrar algunas como la de Alexopoulou (2005) y Arcos Pavón (2009) que analizan los errores morfosintácticos y las de Torijano Pérez (2002), Sánchez Iglesias (2003) y Palapanidi (2010) que se centran en el análisis de los errores léxicos y su comportamiento en el proceso de aprendizaje del español. Entre estos trabajos, es de especial interés para nosotros los de Torijano Pérez (*op. cit.*) y Arcos Pavón (*op. cit.*) porque se dedican a analizar la interlengua de lusohablantes.

Como se ha visto en el capítulo 3, hay una gran variedad de propuestas y de clasificaciones que se han utilizado en los estudios de adquisición de lenguas a la hora de analizar los errores. Dada la especificidad de nuestro contexto, el aprendizaje de una lengua tipológicamente similar, nos interesaba especialmente conocer la tipología del error, es decir, si se daba en el significado (error de tipo formal) o en el significado (error de tipo semántico), y la etiología del error, es decir, si tenía un origen interlingüístico (L1) o intralingüístico (L2). Todo ello nos ha llevado a identificar primeramente los errores según la categoría de palabra, que en la literatura se suele clasificar en palabras léxicas (*lexical words*) y palabras funcionales (*grammatical words*) (Agustín Llach, 2011a), para luego focalizar el análisis en los errores léxicos.

Como se ha visto anteriormente en el marco teórico, se llaman palabras léxicas aquellas que tienen contenido semántico y comprenden los sustantivos, adjetivos, verbos y adverbios. Estas palabras constituyen inventarios abiertos, es decir, como una lengua es un sistema vivo y dinámico, siempre surgen nuevos términos que se van incorporando. Las palabras

funcionales, por otro lado, son aquellas que tienen un significado gramatical, es decir, establecen las relaciones entre las palabras léxicas e incluyen los artículos, pronombres, preposiciones, conjunciones y verbos copulativos. Las palabras funcionales, al contrario de las léxicas, constituyen inventarios cerrados, porque sólo hay un número finito de cada una de ellas.

Al considerar los errores globalmente, se ha observado que su frecuencia disminuye a medida que los estudiantes progresan en el aprendizaje. Esta disminución es estadísticamente significativa entre los niveles intermedio y avanzado. Estos resultados están en línea con las investigaciones de Fernández (1997) y Palapanidi (2010) realizadas en contextos de aprendizaje de ELE por hablantes de lenguas tipológicamente diferentes, en las que también se identificó una disminución en el número de errores en la producción de los estudiantes de nivel avanzado. Por otro lado, los estudios de Hughes y Lascaratou (1982), Santos (1988) y Ellis (2008) también llevados a cabo en contextos de lenguas diferentes tipológicamente demuestran que los errores afectan más a las palabras léxicas y que los errores en esta categoría de palabras se consideran los más graves por hablantes nativos porque interrumpen la comunicación.

Dado que esta investigación se centra en el aprendizaje de español por lusohablantes brasileños, lo que se caracteriza como un contexto de lenguas próximas, nos interesaba averiguar si también en este contexto los errores afectaban más a las palabras léxicas y luego profundizar en las características de los mismos. El análisis de la categoría de palabra afectada por el error reveló que el número de errores en las palabras léxicas es superior al número de errores en las palabras funcionales en los tres niveles, lo que corrobora los resultados de estudios anteriores mencionados más arriba y el interés y la importancia de entender y estudiar los errores léxicos en un contexto de lenguas afines.

Investigaciones dedicadas al análisis del aprendizaje del inglés (L2) reportan resultados en los que la producción escrita de los estudiantes presenta menos errores léxicos a medida que aumenta su competencia lingüística (Wolfe Quintero *et al.*, 1998; Agustín Llach, 2011a). En nuestro caso, se observó que tanto los errores léxicos como los funcionales disminuyen significativamente sólo cuando los participantes alcanzan el nivel avanzado. Esto indica que la disminución de errores tanto léxicos como funcionales en un contexto de lenguas

próximas ocurre solamente tras un mayor número de horas de instrucción y contacto con la lengua meta, lo que corrobora estudios anteriores (ej.: Navés *et al.*, 2005) en los que se ha averiguado que los estudiantes de niveles menos proficientes recurren más a la transferencia como estrategia compensatoria. Con todo, el hecho de que solamente observamos mejora en la frecuencia de los errores en la producción escrita de los participantes de niveles de competencia alto puede deberse a que los alumnos, dada la proximidad tipológica de las lenguas implicadas, recurran a su L1 como estrategia compensatoria durante un período más largo de su aprendizaje, transfiriendo las palabras cognadas equivalentes en la L1 tanto debido a la semejanza de los significantes, como a la disponibilidad de equivalentes traducidos que permiten el acceso de las formas de la L2 a partir de los lemas de la L1. Cuando el producto resultante de la transferencia se corresponde a las normas de la L2, se suele llamar transferencia léxica positiva (Jiang, 2002). Podemos referirnos a este fenómeno para explicar porque sólo tras un mayor número de horas de instrucción y contacto con la lengua meta los alumnos son capaces de demostrar una mayor consciencia de las características propias de la lengua segunda, una vez que son más conscientes de en qué casos pueden o no recurrir a su L1 como estrategia productiva en su intento de escribir en L2.

Además de la fluidez y la diversidad léxica, también se utiliza el error léxico como índice para medir la calidad del texto escrito (Engber, 1995). Así, verificamos que el análisis en conjunto de los resultados de estas tres medidas indica una mejora notable en la calidad de la producción escrita en español por los participantes del estudio tras alcanzar el nivel avanzado. Es decir, es en el nivel avanzado que se observa un mayor dominio de la lengua española, lo que revela una mejora en la producción escrita. De ello, se puede deducir que el número de horas de instrucción y contacto con la lengua meta juega un papel fundamental en el caso del aprendizaje de lenguas próximas como es el del español por lusohablantes tal y como apunta Calvo Capilla y Ridd (2009). Sin embargo, somos conscientes de que otros factores tanto intrínsecos, como la motivación, o extrínsecos, como el tipo de instrucción y el material, también influyen en estos resultados (Shadish *et al.*, 2001), por lo que se tienen que tener en cuenta al generalizarlos a otros contextos.

### 11.3. Los errores según la clase de palabra léxica

**Pregunta 3** ¿A qué clase de palabra léxica (sustantivo, adjetivo, verbo o adverbio) afectan los errores a medida que aumenta el número de horas de instrucción?

**Resultados:** Se han identificado errores en todas las clases de palabra léxica (sustantivos, adjetivos, verbos y adverbios). Sin embargo, en la clase de los verbos se detecta una mayor presencia de errores, seguida de la de los sustantivos. En el caso de los sustantivos, sólo se ha verificado una disminución significativa de los errores entre los niveles inicial y avanzado, mientras que en los verbos, la presencia de los errores disminuye entre los tres niveles. Con respecto a los adjetivos y a los adverbios, los errores son menos frecuentes en el nivel avanzado.

Como se ha visto anteriormente, la fluidez aumenta de manera significativa en el nivel intermedio y se mantiene en el nivel avanzado. Asimismo, este comportamiento se repite en cada una de las clases de palabra léxica (sustantivo, adjetivo, verbo o adverbio) cuando se las analiza por separado. Creemos que nuestros resultados se deben a que en un contexto de lenguas afines como es la de los estudiantes brasileños de español, los alumnos son falsos principiantes porque, como afirma Henriques (2000), el dominio léxico-semántico en su L1 les permite desenvolverse con cierta soltura en la lengua meta dada su proximidad. En el caso de lenguas no tan próximas, el comportamiento puede ser diferente. Por ejemplo, en el estudio de Agustín Llach (2001a) sobre la adquisición del inglés (L2) por alumnos de primaria hablantes de español (L1), los sustantivos se adquieren primero y a medida que los estudiantes progresan en su aprendizaje, la producción de sustantivos disminuye, mientras la de los verbos aumenta.

Hasta aquí y como preámbulo nos hemos referido al uso de las palabras léxicas en general. En este apartado, focalizamos en el estudio de los errores que afectan a las palabras léxicas. Como se ha visto en el capítulo de los resultados (10.2. 'Resultados y análisis de errores léxicos y funcionales'), los errores léxicos considerados globalmente tienden a disminuir a medida que aumenta el número de horas de instrucción, pero sólo lo hacen de manera significativa cuando los participantes alcanzan el nivel avanzado. Para contestar a la tercera pregunta de nuestra investigación, desglosamos los errores léxicos en cuatro categorías correspondientes a las clases de palabra léxica (sustantivos, adjetivos, verbos y adverbios), para ver de qué manera los errores inciden en cada una de ellas. Los resultados han indicado una mayor presencia de errores en los verbos y los sustantivos, aunque con un

comportamiento diferente. Mientras los errores en los verbos disminuyen significativamente a lo largo de los tres niveles, en el caso de los sustantivos no se observa mejora significativa, excepto si comparamos los datos del nivel inicial con los del avanzado. Por otro lado, hemos visto que la presencia de los errores en los adjetivos y adverbios sólo disminuye significativamente en el nivel avanzado. Estos resultados revelan que en un contexto de lenguas próximas el aprendizaje de los adjetivos, adverbios y especialmente de los sustantivos es un proceso lento en el que el desarrollo es gradual y que sólo se puede notar una vez que los estudiantes han alcanzado un nivel de competencia alto, tras un mayor número de horas de instrucción y contacto con la lengua meta.

Con el objetivo de profundizar en la discusión, hemos averiguado el tipo y origen del error en cada una de las clases de palabra léxica, para luego analizar cualitativamente los más frecuentes, ofrecer algunos ejemplos de los errores encontrados en el *corpus* y averiguar si se identifica una tendencia a la fosilización. Precisamos que no es el objetivo del presente trabajo ofrecer listas de errores, sino más bien presentar una visión general de las características de los errores léxicos según el tipo y el origen.

Tras el análisis global de los errores léxicos, se ha constatado que el 74% son de tipo formal. En lo que se refiere a los errores en los verbos y en los sustantivos, en ambos casos el porcentaje de errores formales es el mismo, un 73%. Sin embargo, mientras en el primero el 51% de los errores formales tiene un origen intralingüístico, en el segundo, el 55% son de origen interlingüístico. El análisis de los adjetivos y los adverbios revela que el 72% de los errores en el primer caso, y el 80% en el segundo, son de tipo formal, y que tienen un comportamiento similar al de los verbos, dado que más de la mitad de ellos también tiene un origen intralingüístico (51% en los adjetivos y 63% en los adverbios).

Aunque los errores en los verbos sean más frecuentes que los errores en las demás clases de palabra léxica cuando se considera el *corpus* total, al analizar los datos de cada nivel por separado, vemos que en el nivel avanzado la media de errores en los sustantivos es superior a la de las demás clases de palabra (figura 14). A continuación analizamos algunos ejemplos de los errores más frecuentes en sustantivos y verbos.



Por lo que respecta a errores en sustantivos, los más frecuentes son de tipo formal y tienen su origen en la influencia de la L1 (portugués). Hemos identificado dos tipos de estrategias comunicativas compensatorias predominantes que ejemplificamos<sup>47</sup> a continuación:

- a) Préstamos: transferencias directas de la L1 sin ningún tipo de adaptación a la L2. Al tratarse de la producción lingüística de estudiantes de L2 en un contexto de lenguas próximas, vemos que los errores de préstamos corresponden, muchas veces, a significantes equivalentes o casi idénticos en ambas lenguas, y que el error se debe a pequeños detalles como es el empleo inadecuado del acento gráfico, por ejemplo. Éste es el caso del préstamo ‘*inglês\**’, en la que la única diferencia entre el portugués y el español se da por el uso del acento circunflejo en portugués y la tilde en español (ej.: ‘*inglês*’ (port.)/‘*inglés*’ (esp.)). En otros de préstamo, vemos que el error se debe a la sustitución de consonante por otra, y que esta diferencia identifica la palabra como propia de la L1 o de la L2, como se puede observar en el ejemplo de préstamo ‘*origem\**’ (‘*origem*’ (port.)/‘*origen*’ (esp.)). Con todo, también se dan casos de préstamos en los que los ítems de ambas lenguas difieren considerablemente, como en el ejemplo ‘*milho\**’ (‘*milho*’ (port.)/‘*maíz*’ (esp.)) y ‘*greves\**’ (‘*greves*’ (port.)/‘*huelgas*’ (esp.)). Esta estrategia es muy utilizada en el nivel inicial, pero su frecuencia disminuye a medida que aumenta el número de horas de instrucción. Aún así, se ha detectado una presencia amplia de los préstamos en el nivel avanzado, lo que indica que ésta es una estrategia que se sigue utilizando por estudiantes que tienen un nivel alto de competencia.
- Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>origem</i>	<i>origem*</i>	origen
<i>ensino</i>	<i>ensino*</i>	enseñanza
<i>monstro</i>	<i>monstro*</i>	monstruo
<i>milho</i>	<i>milho*</i>	maíz
<i>inglês</i>	<i>inglês*</i>	inglés
<i>greves</i>	<i>greves*</i>	huelgas
<i>festa</i>	<i>festa*</i>	fiesta

<sup>47</sup> Las palabras con asterisco (\*) indican ejemplos de errores en el *corpus* analizado.

b) Aproximación a la L2. Se ha nombrado así las palabras en las que se observa un esfuerzo por transformar el ítem de la L1 y atribuirle características propias de la L2, dando lugar a formas en las que se fusionan la L1 con características de la L2, como el uso de la consonante 's' en lugar del dígrafo 'ss' propio de la lengua portuguesa para representar el fonema [s]. Otra estrategia bastante utilizada como es el de la sustitución de sufijos de la lengua portuguesa por el correspondiente en la lengua española (ej.: *-são* (port.)/*-sión* (esp.)). Sin embargo, como se ve en el ejemplo abajo, a veces hay que hacer otros ajustes además de la sustitución de los sufijos. Aunque las formas erróneas producidas tengan características de los dos idiomas, consideramos que el error propiamente dicho se debe a los rasgos característicos de la L1, por lo que se han clasificado como de origen interlingüístico. Se han detectado ejemplos que revelan la utilización de estas estrategias a lo largo de todo el período de formación, y el hecho de que estos casos persistan en el nivel avanzado, revela que los estudiantes la emplean ampliamente aún cuando ya han alcanzado un nivel alto de competencia. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>pássaro</i>	pásaro*	pájaro
<i>compreensão</i>	comprensión*	comprensión

El análisis más detallado de los errores nos permitió identificar también ejemplos en los que se observa la utilización de la misma estrategia de producción a lo largo de los tres niveles. Es el caso de la utilización del sufijo *-dad* en lugar del sufijo *-tad*, ambos propios de la lengua española. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>faculdade</i>	facultad*	facultad
<i>dificuldade</i>	dificultad*	dificultad

En estos ejemplos es difícil determinar la causa del error, dado que se observan características de ambas lenguas. Por un lado, se podría inferir que los estudiantes utilizan una estrategia productiva que consiste en eliminar la vocal final del sufijo portugués *-dade* en su intento de aproximarse a la L2 (ej.: *idade* (port.)/*edad* (esp.) y *cidade* (port.)/*ciudad* (esp.)), lo que los caracterizarían como de influencia interlingüística, por otro, puede tratarse de la generalización del uso del sufijo español *-dad* a palabras que no corresponden, por lo que el error sería

intralingüístico. Así, vemos que muchas veces la proximidad tipológica entre las lenguas portuguesa y española impide la detección precisa de la causa del error léxico.

En el *corpus* se observan ejemplos que revelan la presencia de préstamos y de aproximaciones a la L2 en los tres niveles, incluso en el avanzado, lo que indica una tendencia a la persistencia de estas estrategias compensatorias en la interlengua de los estudiantes.

En cuanto a los errores en los verbos, en su mayoría son de tipo formal (errores ortográficos) y gran parte de ellos tiene un origen intralingüístico, entre los cuales hemos identificado los siguientes tipos:

a) Errores de acentuación gráfica:

- Errores por adición del acento gráfico. De acuerdo con la regla ortográfica actual de la lengua española, los monosílabos tónicos no llevan tilde, salvo en el caso de la diferenciación de los monosílabos átonos y tónicos que tienen la misma forma, como en la preposición *de* y la forma verbal *dé*, del verbo *dar*. En el primer ejemplo que presentamos abajo se puede inferir que el estudiante procura distanciarse de su L1, dado que las formas de ambas lenguas son idénticas. Sin embargo, éste no es el caso de los ejemplos siguientes, en los que se ve únicamente una aplicación equivocada de las reglas de acentuación de la L2, según la cual las palabras oxítonas terminadas en vocal deben acentuarse. Consideramos que son errores propios del proceso del aprendizaje, es decir, del desarrollo de la interlengua, que revelan tanto un desconocimiento de las reglas de la L2, como un intento de aproximarse a ella. Detectamos este error en los tres niveles, incluso en el nivel avanzado. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) fui, ir</i>	fuí*	(yo) fui, ir
<i>(ele) foi, ser</i>	fué*	(él) fue, ser
<i>(ele) deu, dar</i>	dió*	(él) dio, dar

- Errores por omisión del acento gráfico. En algunos casos, como el del primer ejemplo que presentamos abajo, correspondiente a la forma verbal de la primera persona del plural de pretérito imperfecto de indicativo del verbo ser, las formas de la L1 y la L2 son equivalentes. En otros, como los ejemplos posteriores correspondientes a las formas verbales de tercera persona del plural del presente de indicativo del verbo español estar y a la tercera persona del plural del futuro de indicativo del verbo español ser, sin embargo, no hay una correspondencia directa entre las dos lenguas. En el primer caso, podemos decir que influye la intuición que tiene el estudiante de que la L2 debe ser diferente de la L1, lo que le lleva a omitir la tilde, aunque según la regla ortográfica de ambos idiomas la acentuación de las palabras esdrújulas (proparoxítonas) es obligatoria. En los siguientes ejemplos, podemos decir que se trata de desconocimiento o aplicación equivocada de las reglas de acentuación de la L2, según la cual las palabras agudas (oxítonas) terminadas en –n deben llevar la tilde. Se ha verificado la presencia de estos errores en los tres niveles, incluso en el nivel avanzado. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(nós) éramos, ser</i>	eramos*	(nosotros) éramos, ser
<i>(eles) estão, estar</i>	estan*	(ellos) están, estar
<i>(eles) serão, ser</i>	seran*	(ellos) serán, ser

El análisis de los errores de acentuación gráfica ha revelado una mayor tendencia a omitir la tilde. En esta misma línea, Durão (2004), en el análisis que realizó de los errores presentes en la interlengua de brasileños que aprenden español, también identificó más errores de omisión que de adición del acento gráfico. En nuestro *corpus*, esta característica predomina en las formas verbales de pretérito imperfecto de indicativo de la segunda y tercera conjugaciones. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) comia, comer</i>	comia*	(yo) comía, comer
<i>(eu) vivia, viver</i>	vivia*	(yo) vivía, vivir
<i>(eu) tinha, ter</i>	tenia*	(yo) tenía, tener
<i>(eu) fazia, fazer</i>	hacia*	(yo) hacía, hacer

Como se ve en los ejemplos, la formación de este tiempo verbal es similar en ambas lenguas, pero difieren en cuanto al empleo de la tilde para señalar la sílaba tónica. Mientras que en portugués no se acentúa la vocal débil cuando forma hiato con la vocal posterior, en español la vocal débil lleva tilde si está en la sílaba tónica para indicar el hiato. En estos errores se observa una fusión de estrategias, que la autora denomina ‘cruce de estrategias’ (Durão, *op. cit.*: 105), en los que actúa por un lado la intuición que tienen los estudiantes de no acentuar este tipo de hiato tal y como indica la regla de su L1 y por otro la aplicación equivocada de la regla de acentuación de la L2 según la que no se acentúan las palabras llanas terminadas en vocal. Así, vemos que estos errores de omisión de la tilde pueden deberse tanto a factores interlingüísticos como intralingüísticos y que dada la proximidad tipológica es difícil determinar exactamente la causa del error. Son muy frecuentes en los niveles inicial e intermedio, pero apenas ocurren en el nivel avanzado.

b) Confusión en la representación gráfica de fonemas:

- Confusión en la grafía del fonema fricativo velar sordo /x/ propio de la lengua española representado gráficamente por las letras ‘g’ y ‘j’. Este error ocurre en casos aislados en textos del nivel avanzado, por lo que no se podría hablar que existe una tendencia general, una vez que probablemente se tratan de errores individuales. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
(eu) <i>pegava, pegar</i>	(yo) <i>cojia*</i>	(yo) <i>cogía, coger</i>

- Confusión en la correspondencia fonográfica del fonema /θ/ y sus alófonos [θ] y [s] al final de sílaba en la primera persona del singular del presente de indicativo de los verbos españoles terminados en –cer. Este error se debe probablemente al hecho de que los estudiantes no hacen la distinción en la pronunciación de la fricativa interdental sorda [θ] propia de la variante peninsular de la lengua española y utilizan la grafía del fonema que corresponde a su pronunciación que se identifica con la pronunciación de la variante del español de América, el sonido

fricativo alveolar sordo [s]. Se ha detectado una mayor presencia de estos errores en el nivel intermedio. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) conheço, conhecer</i>	(yo) conosco*	(yo) conozco, conocer
<i>(eu) agradeço, agradecer</i>	(yo) agradezco*	(yo) agradezco, agradecer

- Confusión en la grafía del fonema fricativo interdental sordo /θ/ y sus alófonos [θ] y [s] representado por las letras 'c' y 'z' ante 'e', que revela desconocimiento de las convenciones que determinan el uso de la 'c' en lugar de la 'z' en este caso. Es posible que esta confusión tenga su origen en la forma del verbo en infinitivo de los verbos españoles terminados en -zar. La presencia de este tipo de error es mayor en los niveles inicial e intermedio, y por el contrario, se detectaron muy pocos ejemplos en el nivel avanzado. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) comecei, começar</i>	(yo) empecé*	(yo) empecé, empezar
<i>(ele) alcance, alcançar</i>	(él) alcance*	(él) alcance, alcanzar
<i>(ele) aprofunde, aprofundar</i>	(él) profundize*	(él) profundice, profundizar

De todos modos, es interesante notar que se observa también esta confusión gráfica en otra forma verbal, como en la primera persona del singular del pretérito indefinido del indicativo del verbo español hacer. Es posible que este error se deba a una generalización de la estrategia anterior o una confusión con la forma de tercera persona del singular hizo. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) fiz, fazer</i>	(yo) hize*	(yo) hice, hacer

c) Confusión en las formas verbales:

- Confusión de formas verbales que ocurre entre la primera y tercera persona del pretérito indefinido de indicativo de verbos irregulares como tener, hacer y poder. Estos errores revelan desconocimiento de las formas de las personas verbales propias de los verbos irregulares. Una posible explicación sobre la causa de estos errores puede deberse a una confusión que tiene origen en las formas regulares del presente de indicativo de los verbos de segunda y tercera conjugación que en este

caso son las mismas tanto en la L1 como en la L2, en las que la desinencia –o indica la primera persona de singular, mientras que la desinencia –e indica la tercera persona de singular (ej.: comer, *eu como* (port.)/*yo como* (esp.), *ele come* (port.)/*él come* (esp.)). Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) tive, ter</i>	(yo) tuvo*	(yo) tuve, tener
<i>(eu) fiz, fazer</i>	(yo) hizo*	(yo) hice, hacer
<i>(eu) pude, poder</i>	(yo) pudo*	(yo) pude, poder

- Confusión de modos verbales que ocurre entre las formas españolas de la tercera persona del plural del pretérito indefinido de indicativo (-ron) y el pretérito imperfecto de subjuntivo (-ran) ocasionada por confusión en la grafía de las desinencias. En este caso, vemos que el error puede tener un doble origen, porque puede deberse tanto a la influencia de la L1, dada la grafía de la desinencia del pretérito indefinido de indicativo (-ram), como a una confusión con la forma existente en la L2, pero que corresponde a otro tiempo y a otro modo verbal, el pretérito imperfecto de subjuntivo. Estos errores ocurren principalmente en el nivel intermedio, probablemente porque es en este nivel que se introduce el contraste de pasados. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eles) foram</i>	(ellos) fueran*	(ellos) fueron
<i>(eles) houveram</i>	(ellos) hubieran*	(ellos) hubieron
<i>(eles) transformaram</i>	(ellos) transformarían*	(ellos) transformaron

#### d) Sobregeneralización:

- Sobregeneralización de la regla de formación del gerundio. Se ha identificado la generalización de la formación del gerundio de las formas regulares a las irregulares. Así, los estudiantes utilizan la raíz del verbo español y añaden la desinencia propia del gerundio de los verbos de segunda y tercera conjugaciones (-iendo), sin darse cuenta que se trata de un verbo irregular que debe sufrir un cambio en la raíz. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>escolhendo (escolher)</i>	elegiendo*	eligiendo, elegir
<i>podendo (poder)</i>	podiendo*	pudiendo, poder

- Sobregeneralización de la diptongación. Los estudiantes perciben la diptongación como una marca propia de la L2 y por ello extienden el empleo de esta característica a formas que no la aceptan en el intento de producir L2 y alejarse de su L1. La forma errónea ‘espiero\*’, es un ejemplo de ello, dado que las formas de la L1 y la L2 son idénticas (ej.: *espero* (port.)/ *espero* (esp.)). Otras veces, los estudiantes emplean la irregularidad verbal representada por la diptongación en casos en los que no se aplican, como en el ejemplo ‘tiengo\*’, probablemente por analogía a las formas *tienes, tiene, tienen*. También se ve un caso similar en el ejemplo ‘empiecé\*’, en el que los estudiantes extienden una irregularidad propia del presente de indicativo (*empiezo, empiezas, empieza, empiezan*) a las formas de pretérito indefinido de indicativo (*empecé*). Estos errores son más frecuentes en el nivel inicial, pero persisten en los niveles intermedio y avanzado. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) espero, esperar</i>	(yo) espiero*	(yo) espero, esperar
<i>(eu) tenho, ter</i>	(yo) tiengo*	(yo) tengo, tener
<i>(eu) comecei, começar</i>	(yo) empiecé*	(yo) empecé, empezar

- e) Hipercorrección. Algunas veces, cuando la grafía de la palabra es similar en ambas lenguas, los estudiantes optan por diferenciarlas, adaptándola a lo que cree que se corresponde a la L2. Es el caso de las formas erróneas ‘tube\*’, ‘tubo\*’ y ‘obtube\*’ del pretérito indefinido de indicativo, que se escriben con *v* tanto en L1 como en L2. También puede deberse a una analogía con la forma española *hube*, del verbo *haber*. Además, se debe tener en cuenta que en este error también está implicado otro factor que es la confusión en la representación gráfica del fonema oclusivo bilabial sordo /b/ del español que puede darse a través de las letras ‘v’ y ‘b’. Sólo se ha identificado este error a partir del nivel intermedio y su presencia es más abundante en el nivel avanzado. El hecho de que en el nivel inicial no se ha detectado este error, se debe probablemente porque es a partir del nivel intermedio que se estudian las formas irregulares del pretérito indefinido de indicativo. Ejemplo:



Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) tive, ter</i>	(yo) tube*	(yo) tuve, tener
<i>(ele) teve, ter</i>	(él) tubo*	(él) tuvo, tener
<i>(eu) obtive, obter</i>	(yo) obtube*	(yo) obtuve, obtener

También hemos analizado cualitativamente los errores formales de origen interlingüístico en los verbos, lo que ha revelado que la confusión ortográfica tiene un peso considerable en la frecuencia de este tipo de error. A modo de ejemplo, destacamos la confusión entre las representaciones gráficas del fonema /b/ que se observa en la escritura de las formas de pretérito imperfecto de indicativo que en portugués se escribe con 'v' (ej.: *cantava*), mientras que en español se escribe con 'b' (ej.: *cantaba*). Esta confusión lleva a veces a que el estudiante produzca formas que son idénticas a las de su L1 debido a la semejanza lingüística. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) estava, estar</i>	(yo) estava*	(yo) estaba, estar
<i>(eu) buscava, buscar</i>	(yo) buscava*	(él) buscaba, buscar
<i>(eu) tomava, tomar</i>	(yo) tomava*	(yo) tomaba, tomar

En otras ocasiones, las formas producidas reflejan una fusión de las dos lenguas, porque utilizan la raíz de verbos propios de la lengua española, pero la desinencia tiene rasgos de la L1, lo que genera formas que no hacen parte de ninguno de los dos sistemas lingüísticos implicados. De todos modos, aunque la forma errónea producida tenga características de las dos lenguas, consideramos que el error se debe al empleo de la desinencia verbal propia de la L1. Este error se ha detectado principalmente en los niveles iniciales, aunque se ha identificado la confusión en la grafía del par 'v' y 'b' en menor número también en textos de los niveles avanzados. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>(eu) chegava, chegar</i>	(yo) llegava*	(yo) llegaba, llegar
<i>(nós) falávamos, falar</i>	(nosotros) hablavamos*	(nosotros) hablábamos, hablar
<i>(eles) escutavam, escutar</i>	(ellos) escuchavan*	(ellos) escuchaban, escuchar

Como hemos visto en los resultados del análisis cuantitativo, los errores en los verbos disminuyen significativamente a lo largo del proceso de aprendizaje, pero el análisis cualitativo de los errores más representativos indica que en el nivel avanzado se siguen

detectando errores formales como la omisión/inclusión del acento gráfico y confusión en la grafía de algunos fonemas. Aunque se trate de una convención, los errores en la ortografía reflejan además de desconocimiento de reglas básicas de la L2, algunas estrategias de producción y dificultades propias de los estudiantes brasileños cuando tienen que escribir en español. Consideramos que tanto la proximidad tipológica como el hecho de que estos errores no se consideran graves, puesto que no suelen interrumpir la transmisión del mensaje, son factores que contribuyen a su permanencia en la interlengua de los estudiantes con un nivel de competencia alto, incluso en un contexto de formación de profesores de ELE en el que se espera que los participantes tengan un dominio de las reglas del sistema de la L2.

#### 11.4. Los errores léxicos según el tipo

---

**Pregunta 4** ¿Qué tipo de errores léxicos (formal, semántico o mixto) se observan en los textos escritos por profesores brasileños de ELE en formación según el número de horas de instrucción?

**Resultados:** Los errores formales son mucho más numerosos que los errores semánticos o mixtos. Sin embargo, aunque los errores formales disminuyen a medida que aumenta el número de horas de instrucción, ellos siguen siendo los más abundantes en los textos del nivel avanzado. Los errores semánticos son mucho menos frecuentes que los errores formales, pero no se observa mejora entre los niveles. Los errores mixtos son muy poco frecuentes y disminuyen significativamente cuando los participantes alcanzan el nivel avanzado.

El objetivo del análisis de los errores según el tipo en las palabras léxicas es poder identificar si las dificultades de los estudiantes se encuentran en el significante, es decir, si la forma de las palabras utilizadas por los alumnos corresponden a la normativa ortográfica vigente de la lengua española (L2), o si la confusión se da en el ámbito del significado de las palabras.

Esta clasificación de los errores apoyada en la distinción entre el aspecto formal y semántico (James, 1998) se basa en la constatación de que el lexicón mental se organiza por asociaciones semánticas o formales tanto en la L1 como en la L2, según el modelo de organización léxica que se basa en la distinción entre comprensión y producción lingüística (López Morales, 1993). Con todo, Singleton (1987) afirma que en la organización del lexicón mental también influyen factores como la manera en que el estudiante adquirió la palabra,

el grado de conocimiento que tiene de la palabra y la percepción que tiene de la semejanza formal o semántica entre la palabra de la L2 y la de su L1.

Como esta investigación se dedica al análisis de la interlengua que es el producto del proceso de aprendizaje de español (L2) por lusohablantes brasileños, es decir, un contexto de lenguas afines, tendemos a pensar que la organización del lexicón mental de la L2 (español) se asemeja y se relaciona en gran medida con la organización del lexicón mental de la L1 (portugués). Sin embargo, aunque investigaciones como la nuestra colaboren para la construcción del conocimiento que se tiene sobre la adquisición del vocabulario y el aprendizaje de segundas lenguas en un contexto de lenguas próximas, es necesario seguir investigando para profundizar en ello y averiguar hasta qué punto la organización del lexicón mental de la L2 se asemeja al de la L1.

En el *corpus* analizado se observan errores léxicos de tipo formal, semántico y también mixto. Tal y como se ha explicado anteriormente, identificamos como errores formales aquéllos que afectan a la representación gráfica de la palabra cuando ésta no se corresponde a las normas ortográficas vigentes de la lengua española (L2). Los errores semánticos son aquéllos que afectan al significado de la palabra y los identificamos en los ítems cuyo significado en español (L2) no se adecua al contexto en el que está insertado. Cuando identificamos en una palabra errores que afectan tanto a la forma como al significado, los clasificamos como errores mixtos.

Como se ha visto en el capítulo de los resultados, los **errores formales** son los más frecuentes en la muestra en comparación con los semánticos y mixtos, pero disminuyen significativamente a lo largo del período de formación. Estos resultados revelan que, tal y como se ha visto en otros estudios (ej.: Celaya y Torras, 2001; Woodall, 2002; Navés *et al.*, 2005; Celaya, 2007), el número de horas de instrucción incide positivamente en la superación de los errores léxicos de forma también en un contexto de aprendizaje de lenguas próximas. Con todo, consideramos que la media de errores de tipo formal por texto en la producción escrita de estudiantes con un nivel de competencia alto sigue siendo bastante elevada (M=5,16) si la comparamos con la media de los errores semánticos (M=1,19) y mixtos (M=0,75). Estos datos nos han llevado a realizar un análisis más detallado de los errores formales para averiguar si son errores que también se dan en los niveles

anteriores, y todavía no se han superado, y consecuentemente se verifica una tendencia a la persistencia de este tipo de error en la interlengua de estos estudiantes.

Al examinar los errores formales de forma cualitativa, observamos que muchos de esos errores se deben básicamente a cuatro factores: a) al desconocimiento de las reglas de representación ortográfica del español (L2); b) a la transferencia de las reglas y formas de la L1; c) a la aplicación de la estrategia de sobregeneralización de algunas de las reglas de la L2, o incluso d) a la hipercorrección, estrategia que los estudiantes aplican cuando creen que determinado ítem en su producción tiene que ser distinto de su L1, e intenta alejarse de ella y aproximarse a la L2, alterando palabras de manera innecesaria en algunos casos y creando formas que no se corresponden ni a la L1 ni a la L2.

Como se ha explicado en la discusión de las preguntas anteriores, exponemos en este apartado, a modo de ejemplo, algunos de los errores formales más abundantes identificados en el *corpus* de esta investigación. Se puede pensar que un análisis de los errores formales no es tan importante porque este tipo de error no suele afectar negativamente la comunicación y que tarde o temprano se superarán. Con todo, destacamos la importancia del análisis de los errores formales en un contexto de lenguas próximas como son el portugués (L1) y el español (L2), especialmente en un curso de formación de profesores, porque, como se verá más adelante en los ejemplos, por un lado la diferencia entre escribir en portugués o escribir en español está muchas veces en algo tan sutil como en la correcta utilización del acento gráfico, y por otro, porque los profesores deben tener un dominio de la ortografía de la lengua española para ser capaces de ofrecer una muestra de lengua libre de errores y para guiar a sus alumnos de manera adecuada en su proceso de aprendizaje del español.

Así, con relación a la acentuación gráfica, podemos agrupar los errores bajo los siguientes conceptos:

- a) errores de omisión: hemos encontrado una diversidad de errores de omisión del acento gráfico, como se puede observar a partir de los ejemplos que enseñamos a continuación.

- Errores de omisión del acento gráfico en palabras idénticas en portugués (L1) y en español (L2). Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>país</i>	pais*	país
<i>próximo</i>	proximo*	próximo

Las palabras '*país*' y '*país*' (L1, portugués) corresponden a '*padres*' y '*país*' (L2, español), respectivamente. Es por ello que la omisión del acento gráfico puede conllevar también confusión de sentido en algunos casos, aunque el contexto ayuda a aclararlo.

- Omisión del acento gráfico en palabras similares, en las que hay correspondencia en la sílaba tónica, aunque se marquen con acentos gráficos distintos en cada una de las lenguas. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>terão</i>	tendran*	tendrán
<i>avião</i>	avion*	avión
<i>acadêmica</i>	academica*	académica
<i>também</i>	tambien*	también

- Omisión del acento gráfico cuando las reglas de la L1 y L2 difieren. En los ejemplos abajo las formas de la interlengua coinciden con la regla de acentuación de la L1, lo que ocasiona el error. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>aprendi</i>	aprendi*	aprendí
<i>conheci</i>	conoci*	conocí
<i>depois</i>	despues*	después

En el primer ejemplo, la única diferencia entre las formas de las dos lenguas se da en el empleo de la tilde, por lo que se podría tratar de un préstamo, mientras que en los demás, el significante utilizado corresponde a los de la L2, con excepción de la acentuación gráfica, por lo que se puede decir que el estudiante no recurre a su L1, ya que demuestra conocer la forma correspondiente a la lengua meta, aunque de manera incompleta.

b) Errores de adición: también hemos encontrado errores que se deben al empleo innecesario de la tilde.

- Adición del acento gráfico que coincide con las reglas de la L1, pero difieren de las de la L2. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>sítio</i>	sítio*	sitio
<i>médio</i>	médio*	medio
<i>ódio</i>	ódio*	odio
<i>próprios</i>	próprios*	proprios

En los tres primeros ejemplos, los significantes son idénticos en ambas lenguas, con excepción del empleo de la tilde. Esto dificulta la clasificación precisa del error, porque no podemos decir con exactitud si se trata de un préstamo o simplemente de un error de adición del acento gráfico. En el último ejemplo, con todo, los significantes no se corresponden totalmente en ambas lenguas, lo que revela desconocimiento de la forma propia de la L2. Por lo tanto, en este caso, podemos afirmar que el error se trata de préstamos.

- Adición del acento gráfico en palabras que no lo requieren en ninguna de las dos lenguas. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>penais</i>	penáiles*	penales
<i>encantam</i>	encántan*	encantan

En estos dos ejemplos, vemos que el estudiante identifica correctamente la sílaba tónica, pero la señala innecesariamente. Podría tratarse de un ejemplo de hipercorrección.

c) Errores de sustitución: se ha notado la utilización de signos ortográficos que no pertenecen al sistema ortográfico de la lengua española, como el acento circunflejo (^), la tilde nasal (~) sobre la vocal a (ã) o la ce cedilla (ç). Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>inglês</i>	inglês*	inglés
<i>português</i>	português*	portugués
<i>violência</i>	violência*	violencia
<i>não</i>	não*	no
<i>aversão</i>	aversão*	aversión
<i>cidadã</i>	cidadã*	ciudadana
<i>forçados</i>	forçados*	forzados
<i>dançar</i>	dançar*	danzar/bailar
<i>segurança</i>	segurança*	seguridad

En lo que se refiere al empleo del acento circunflejo (^), vemos que los errores pueden deberse tanto a un despiste en la utilización de la tilde como a préstamos. En el caso de la tilde (~), vemos claramente que se trata de préstamos, porque en los ítems seleccionados, los significantes de la L1 y la L2 no son idénticos. Finalmente, con relación a la cedilla (ç), los dos primeros ejemplos no permiten entrever si se tratan de préstamos o error ortográfico dada la semejanza que se observa entre los ítems de ambas lenguas, mientras que el último no deja lugar a dudas de que se trata de un préstamo directo de la L1, dado que los ítems de la L1 y la L2 son distintos.

Aparte de los errores de acentuación gráfica, también se ha encontrado una variedad de errores de adición, omisión o sustitución de signos ortográficos que presentamos en los ejemplos abajo:

- a) Error de omisión. Este error ocurre cuando se omite una de las letras que forman los encuentros consonánticos propios de la L2:
- El encuentro consonántico 'cc'. En este caso, se puede observar que los significantes de la L1 y la L2 difieren, y que los ejemplos de interlengua revelan cierto conocimiento de las formas de la L2, aunque incompleto.

Portugués (L1)	Error (interlengua)	Español (L2)
<i>construção</i>	construcción*	construcción
<i>interação</i>	interacción*	interacción

- El encuentro consonántico ‘ct’. En este caso, observamos que, a diferencia del caso anterior, los ejemplos de interlengua coinciden con las formas de la L1, por lo que los clasificamos como préstamos.

Portugués (L1)	Error (interlengua)	Español (L2)
<i>estruturas</i>	estruturas*	estructuras
<i>contatos</i>	contatos*	contactos

b) Error de adición. Se verifica este error cuando la palabra analizada presenta un elemento extra, que no corresponde a la forma de la L2.

- Se ha detectado algunos casos en los que los errores se deben a la adición equivocada de una letra. Un ejemplo de ello es el de la inclusión de la vocal ‘e’ en formas derivadas de los verbos ‘ver’ y ‘leer’ cuando no corresponde, probablemente debido a una asociación inadecuada con la forma del infinitivo del verbo ‘leer’, lo que indica el empleo de la estrategia de hipercorrección en el intento de alejarse de la L1. Ejemplos:

	Portugués (L1)	Error (interlengua)	Español (L2)
• L	<i>(eu) leio, ler</i>	leeo*	(yo) leo, ler
a	<i>ver</i>	veer*	ver

- Diptongación. Ésta es una característica que los hablantes brasileños asocian como propia del español (L2). Así, debido a la semejanza lingüística, los estudiantes introducen diptongos en las palabras de la L1 en su intento de producir L2, aún cuando los significantes en ambas lenguas son idénticos. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>rapidamente</i>	rapidamente*	rapidamente
<i>momento</i>	momiento*	momento
<i>atendidas</i>	atiendidas*	atendidas

Esta estrategia compensatoria es una de las características principales del fenómeno conocido como el ‘portuñol’ y se sigue aplicando en los niveles avanzados, aunque en menores proporciones.



c) Error de sustitución (confusión de grafías):

- Confusión en la grafía del fonema /b/. Como se ha visto en el apartado anterior (11.3. Los errores según la clase de palabra léxica), las formas de pretérito imperfecto de indicativo son una fuente productiva de errores, ocasionados tanto por el hecho de que en lengua española (L2) un único fonema /b/ se pueda representar con dos grafías ('b' y 'v'), como por el hecho de que en la lengua portuguesa (L1) la grafía de la desinencia de esta forma verbal se hace con la letra 'v'. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>necessitava</i>	necessitava*	necesitaba
<i>devorava</i>	devorava*	deboraba

Como se puede ver en el primer ejemplo, la forma de la interlengua se identifica con la de la L1, por lo que podemos hablar de préstamo con seguridad. Sin embargo, en el segundo caso, como los significantes de la L1 y la L2 difieren sólo por una letra, no sabemos con exactitud si se trata de una confusión en la correspondencia fonográfica o si es un préstamo de su L1, el portugués.

- Confusión en la grafía del fonema /k/. Esta confusión se observó principalmente delante del encuentro vocálico 'ua', seguramente debido la interferencia de la L1, por lo que vemos una vez más ejemplos de préstamos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>quando</i>	quando*	cuando
<i>quarto</i>	quarto*	cuarto

- Confusión en la grafía del fonema /s/. Se identificó un amplio uso de la doble 'ss' (dígrafo) para representar este fonema, ciertamente debido a la influencia de la L1, resultando en ejemplos de formas híbridas:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>passei</i>	passé*	pasé
<i>associações</i>	asociaciones*	asociaciones

Se nota en estos ejemplos que pese a que utilicen el dígrafo 'ss' propio de la L1, el significante se aproxima más a las características de la L2.

- Confusión en la grafía del sonido /x/. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>pegava</i>	cojia*	cogía
<i>estrangeira</i>	estrangera*	extranjera

En el primer ejemplo, es más probable que se trate de un error de confusión en la representación del fonema /x/, dado que no existe en portugués y que se puede hacerlo tanto con la letra 'g' como con la 'j' en español. Además, dado que la forma 'coger' no existe en portugués, se observa que la interlengua del estudiante presenta indicios de que ha superado el estadio inicial, una vez que ya no recurre a la forma de su L1 y emplea una palabra propia de la L2.

- Confusión en la grafía de sonidos nasales. En portugués (L1), para la representación gráfica de los sonidos nasales a final de sílaba se emplea generalmente la letra 'm', mientras que en español (L2) se suele hacer con la letra 'n'. Los estudiantes son conscientes de ello y utilizan este conocimiento como una estrategia, adaptando las palabras de la L1 cuando quieren producir L2. Sin embargo, como se puede observar en los ejemplos abajo, a veces esta estrategia conduce al error. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>assim</i>	assin*	así
<i>bom</i>	bon*	bueno

- Grafía del sonido nasal palatal sonoro /ñ/. En portugués (L1), este fonema se representa gráficamente a través del dígrafo 'nh' mientras que en español (L2) se emplea la letra 'ñ'. Se han identificado casos en los que los estudiantes adaptaron la grafía del significante de la L1 en su intento de producir L2, lo que resultó en ejemplos creativos, pero que desafortunadamente no corresponden a la L2, como se puede observar a continuación:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>dinheiro</i>	diñero*	dinero
<i>boneca</i>	buñeca*	muñeca

En el primer ejemplo, el equivalente en L2 no contiene el fonema /ñ/, mientras que en el segundo, dado que los ítems de ambas lenguas tienen lexemas distintos, el empleo de esta estrategia no ha funcionado plenamente.

Tras la exposición y explicación de algunos ejemplos de errores identificados en el *corpus*, destacamos que aunque la frecuencia de estos errores disminuye en la producción escrita del nivel avanzado, seguimos identificando errores formales que consideramos que ya debían haberse superado en un nivel de competencia alto. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>assado</i>	assado*	asado
<i>cozido</i>	cozido*	cocido
<i>produzida</i>	produzida*	producida
<i>raízes</i>	raízes*	raíces
<i>estrangeiros</i>	extrangeros*	extrangeros
<i>reação</i>	reaccion*	reacción
<i>espetáculo</i>	espetáculo*	espectáculo
<i>contatos</i>	contatos*	contactos
<i>pretendo</i>	pretiendo*	pretendo
<i>família</i>	família*	familia
<i>dia</i>	dia*	día
<i>jovem</i>	jóven*	joven
<i>saúde</i>	salúd*	salud
<i>governo</i>	gobierno*	gobierno
<i>forçados</i>	forçados*	forzoso
<i>buscava</i>	buscava*	buscaba
<i>(eu) tive, ter</i>	tube*	(yo) tuve, tener

Como se puede observar, el análisis de los errores formales presentes en la producción escrita del nivel avanzado revelan que, por un lado los estudiantes siguen apoyándose en estrategias de transferencia de la L1, de generalización de las reglas de la L2 y de hipercorrección, y por otro, siguen demostrando que desconocen reglas básicas del ámbito ortográfico de la L2. Así, podemos afirmar que tras los análisis de los errores formales, se ha

identificado una tendencia a la persistencia de los mismos en la interlengua de los estudiantes.

En lo que se refiere a los **errores léxicos de tipo semántico**, su frecuencia en el *corpus* analizado ha sido muy inferior a la de los errores formales. Asimismo, no se observa diferencia estadísticamente significativa entre los niveles. Consideramos que la identificación de los errores semánticos en nuestra muestra, un contexto de lenguas próximas, es una tarea difícil. Esta dificultad reside en que algunos de esos errores se dan en palabras que existen en ambas lenguas (portugués L1 y español L2) y que tienen significados similares, pero su frecuencia de uso en la L2 es muy baja o la palabra se restringe a un uso más bien literario. Esta inadecuación en el uso del vocabulario se debe a la frecuencia de uso o a la conveniencia pragmática y, según Calvi (2004), es muy común en el ámbito del aprendizaje de lenguas afines como en el caso del aprendizaje de español por hablantes de italiano. Con todo, pensamos que la baja ocurrencia de los errores semánticos se puede deber a que los alumnos, ante la duda, empleen la estrategia de la evitación (Laufer, 2000).

Una de las causas probables del error semántico puede deberse a la utilización de diccionarios bilingües que suelen ofrecer al estudiante solamente el término correspondiente en la L2, sin matizar las implicaciones que tiene la utilización de cada ítem en diferentes contextos. Es por ello que, en el ámbito de la enseñanza de lenguas afines, como en el del aprendizaje de español por lusohablantes, creemos que es importante la utilización de diccionarios monolingües de español desde los niveles iniciales, dado que los alumnos son capaces de comprender las explicaciones debido a la semejanza lingüística existente entre las lenguas.

Asimismo, la baja frecuencia de los errores de tipo semántico que se ha observado en el *corpus* y el hecho de no haber encontrado diferencia significativa entre los tres niveles son resultados que nos han sorprendido, porque esperábamos que fueran más frecuentes en los niveles iniciales y que se observara una disminución significativa en la frecuencia de este tipo de error en algún momento del período de formación. Una posible explicación para la baja ocurrencia de este tipo de error podría deberse a la proximidad de las lenguas implicadas, dado que son lenguas romances que comparten un alto porcentaje del léxico. Así, la utilización adecuada de este conocimiento previo aportado por L1 y aplicado a la

producción en L2 se ve reflejada en la escasa presencia de los errores semánticos. Otro factor que puede haber incidido en la baja presencia de errores semánticos se debe al contenido de los temas propuestos para las redacciones, pues se trataban de temas cercanos a los participantes que tenían el objetivo de estimular la producción escrita, por lo que no supusieron ningún obstáculo para la comunicación del mensaje.

Por lo tanto, creemos que los temas propuestos fomentaron el uso de un vocabulario general, propio de la esfera personal y familiar, que no exigió el uso de palabras más específicas ni sofisticadas. Para profundizar en el estudio de los errores léxicos de tipo semántico en un contexto de lenguas próximas creemos que sería necesario utilizar pruebas diseñadas específicamente para este fin. Sólo así se podría comprobar si la tendencia observada en la presente investigación se confirma o no. Igualmente, el comportamiento de los errores léxicos observado en la presente investigación revela que son errores que tienen una tendencia a estabilizarse en la interlengua de los estudiantes, dada la dificultad de detectarlos y consecuentemente tratarlos. Con el objetivo de elucidar esta cuestión, analizamos cualitativamente los errores léxicos de tipo semántico presentes a lo largo de los tres niveles para averiguar si son los mismos o al menos si son errores fruto de la misma estrategia de producción.

En el análisis cualitativo de los errores léxicos de tipo semántico hemos encontrado:

- a) Semejanza fonológica. Hay algunos errores semánticos originados por semejanzas fonológicas entre la L1 y la L2, como en el caso del error arreglar\*, probablemente por asociación con *arranjar*. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>arranjar um emprego</i>	arreglar* un empleo*	conseguir un trabajo

- b) Falsos amigos. También se han visto casos en los que el error puede deberse a la confusión que originan los falsos cognados, cuando el significante existe en ambas lenguas, pero tiene significados diferentes en cada una de ellas, como lo advierte Larrinaga (2006) cuando afirma que la similitud formal induce al estudiante a creer que hay también una equivalencia de significados. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>cidade graciosa</i>	ciudad <b>graciosa*</b>	ciudad elegante/hermosa

- c) Colocación. Algunas palabras existen en ambas lenguas con significado similar, pero no se utilizan en los mismos contextos. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>tive vontade de estudar</i>	siempre tuve <b>voluntad*</b> de estudiar español	tuve ganas de estudiar

Polisemia. En algunos casos, un significante en L1 tiene varias acepciones que corresponden a significantes diferentes en la L2. Éste es el caso del verbo ‘ficar’ en portugués. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>As disputas fizeram com que o clima ficasse difícil</i>	Las disputas hicieron con que el clima <b>se quedase*</b> difícil	Las disputas crearon un ambiente hostil
<i>As cidades de Cabo y Palos de la Frontera ficaram amigas</i>	Cabo <b>se quedó*</b> ciudad amiga de Palos de la Frontera	Cabo pasó a ser/ fue nombrada ciudad amiga de Palos de la Frontera

En ambos ejemplos vemos que los estudiantes optan por traducir el verbo ‘ficar’ como ‘quedarse’. Al analizar el segundo ejemplo a fondo, “Cabo se quedó\* ciudad amiga”, observamos que, en su texto, el estudiante se refería al hecho de que la ciudad del Cabo de Santo Agostinho en Pernambuco, Brasil, más conocida como Cabo, había sido nombrada ciudad amiga de Palos de la Frontera, en Andalucía, porque Vicente Yáñez Pinzón, navegador español originario de esta ciudad, había llegado a Cabo de Santo Agostinho en enero del año 1.500. Se trata por lo tanto de un error semántico y su posible causa puede deberse al carácter polisémico del verbo portugués ‘ficar’, que en español puede corresponderse tanto al verbo “estar” (localización: ¿Dónde está el cine? (esp.)/*Onde fica o cinema?* (port.)), como a una variedad de verbos, a los que se suele llamar ‘verbos de cambio’ (Milani, 1999): ponerse, volverse, hacerse, quedarse, convertirse. Así, al expresar la idea de que “as cidades ficaram amigas” (portugués, L1), es decir, las ciudades se convirtieron en ciudades amigas, el alumno optó por utilizar ‘quedarse’. Creemos que eso se debe a que la correspondencia entre ‘ficar’ y ‘quedarse’ se establece al principio de su aprendizaje, y que más adelante, en un nivel de competencia intermedio, cuando se

introducen los verbos de cambio, los estudiantes demuestran tener dificultad en utilizar estos verbos para expresar los diferentes matices que conlleva cada uno de ellos en su expresión lingüística.

Encontramos ejemplos similares a los que presentamos en las producciones escritas en los tres niveles, incluso en el nivel avanzado. Por lo tanto, podemos decir que aunque el análisis cuantitativo revele una frecuencia muy baja de los errores léxicos semánticos y no se haya verificado diferencia estadística significativa entre los niveles, verificamos que esos errores están presentes a lo largo de todo el período de formación y que por esta razón son susceptibles de persistir en la interlengua de los participantes.

Finalmente, discutimos los resultados del análisis de los **errores mixtos**, aquéllos que afectan a la forma y al significado, los menos frecuentes en el *corpus*. La frecuencia de estos errores disminuye de manera significativa entre los niveles intermedio y avanzado. El análisis más detallado de estos errores revela algunas de las estrategias compensatorias de producción utilizadas por los estudiantes. Comentamos una de ellas a partir del siguiente ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>saudade</i>	Recuerdo este tiempo con mucha <b>saludad*</b>	nostalgia

La palabra ‘saludad\*’ identificada en el *corpus* no existe ni en portugués (L1) ni en español (L2), un ejemplo de acuñación en el que influyen ambas lenguas. Si no se tiene en cuenta el contexto, se podría pensar que se trata de la palabra española ‘salud’ o incluso ‘saludar’, escrita de manera incorrecta, dada la similitud formal. Sin embargo, el contexto revela que se trata en realidad de la palabra portuguesa ‘*saudade*’ adaptada a la forma de lo que el estudiante percibe como la equivalente en lengua española (L2). Aunque estos errores sean muy interesantes para el investigador, la baja frecuencia en la muestra y la disminución significativa en el nivel avanzado no indican tendencia a la persistencia de este tipo de error en la interlengua de los estudiantes, aunque sí se puede fosilizar en la interlengua del individuo.

Varios estudios en ASL averiguaron que en los niveles iniciales la producción de los estudiantes presenta más errores de tipo formal, y que a medida que ellos progresan en el aprendizaje, los errores formales van dando lugar a errores semánticos (ej.: Fernández,

1997; Celaya y Torras, 2001; Navés *et al.*, 2005; Palapanidi, 2010). Con todo, nuestros resultados sólo coinciden en parte con los mencionados anteriormente. En la presente investigación, se ha visto igualmente que los errores formales son los más frecuentes en los niveles iniciales y que disminuyen significativamente a medida que los estudiantes progresan en su aprendizaje, pero, al contrario de lo que encontraron los estudios previos, los errores formales siguen siendo los más numerosos en el nivel avanzado. Tal y como se ha dicho anteriormente, creemos que este hallazgo en esta investigación está relacionado con la proximidad lingüística de las lenguas portuguesa (L1) y española (L2) y es un indicio de que es necesario llamar la atención del estudiante para las reglas ortográficas a través de una enseñanza contrastiva para que los futuros profesores de ELE sean capaces no sólo de superar estos errores, sino también de transmitir este conocimiento a sus alumnos.

### 11.5. Los errores léxicos según el origen

---

**Pregunta 5** ¿Cuál es el origen de los errores léxicos en la producción escrita de estos estudiantes a lo largo de los tres niveles? ¿Los errores se deben a la influencia de su L1 (error interlingüístico), el portugués variedad brasileña, o son fruto del proceso de desarrollo de la interlengua (error intralingüístico)?

**Resultados:** Se ha verificado la ocurrencia tanto de errores interlingüísticos como intralingüísticos e incluso ambiguos. También se ha detectado la disminución de la frecuencia de estas tres categorías de errores a lo largo del proceso de aprendizaje. En el primer caso, la mejora es estadísticamente significativa entre los tres niveles de aprendizaje, mientras que en el segundo y el tercero, esta mejora se observa solamente entre los niveles intermedio y avanzado.

El estudio del origen del error en las palabras léxicas tiene el objetivo de identificar si la causa de los errores se debe a la utilización de la estrategia de transferencia de aspectos de la primera lengua o lengua materna (L1) del hablante o a factores propios del proceso de aprendizaje y desarrollo de la interlengua, como son la sobregeneralización y la hipercorrección, por ejemplo. Como se ha visto en el capítulo 3 dedicado al error, los investigadores suelen utilizar metodologías de análisis de errores *ad hoc*, es decir, específicamente diseñadas para atender las características de su propio estudio, lo que ha generado una gran variedad de clasificaciones. Sin embargo, el artículo de Penadés Martínez (2003) sobre las clasificaciones de los errores en el ámbito de la enseñanza de español muestra que el estudio del origen de los errores es un tema recurrente en la literatura y que



hay correspondencia entre varios de los criterios utilizados con diferentes nomenclaturas, como se puede ver en la correlación que la autora establece específicamente sobre éstos: criterio etiológico (Vázquez, 1991), criterio etiológico-lingüístico y gramatical (Santos Gargallo, 1993) y criterio de la descripción lingüística (Fernández, 1997).

Identificamos, en el *corpus* de análisis, errores con un origen tanto en el portugués, la L1 de los participantes (error interlingüístico), como en el propio proceso de aprendizaje de la L2, el español (error intralingüístico). Además, dado que no siempre es fácil identificar el origen del error, menos todavía en situaciones de lenguas afines como en este caso, se identificaron algunos errores como ambiguos (Dulay *et. al.*, 1982), es decir, errores que tienen su origen tanto en la L1 de los estudiantes como en el proceso de aprendizaje de la L2.

Así, constatamos que los errores léxicos identificados en la muestra se reparten entre los de origen **interlingüístico** (49,79%), los errores de origen **intralingüístico** (45,03%) que se restringen a un porcentaje algo inferior, y los errores **ambiguos** (5,18%), que son muy poco frecuentes en comparación con los anteriores. Estos resultados contradicen estudios previos sobre el aprendizaje de ELE por hablantes de L1 diferentes tipológicamente en los que se había puesto de manifiesto que los errores por interferencia de la L1 eran menos frecuentes que los errores intralingüísticos (ej.: Fernández, 1997; Palapanidi, 2010). Esta diferencia de resultados puede deberse a la relación de proximidad de las lenguas implicadas en el contexto de la presente investigación, el portugués (L1) y el español (L2), que como se ha dicho anteriormente, comparten una gran parte del léxico (Israel, 2001).

A partir de la concepción del aprendizaje de una lengua como un proceso, nos interesaba averiguar también si se daban cambios con respecto al origen de los errores a lo largo del período de formación en un contexto de lenguas próximas. El análisis del origen de los errores en cada uno de los tres niveles de competencia ha revelado que los **errores interlingüísticos** disminuyen significativamente entre los tres niveles analizados, lo que sigue los presupuestos reflejados en la literatura de que la influencia de la L1 disminuye a medida que los estudiantes progresan en su aprendizaje (Agustín Llach, 2010). Con todo, también hemos identificado que pese a la disminución, éstos son los errores más numerosos tanto en el nivel inicial como en el avanzado, con relación a los errores intralingüísticos y

mixtos. Este dato revela que los estudiantes con un nivel de competencia alto siguen recurriendo a su L1 cuando intentan producir textos escritos en la lengua meta, tal y como lo hacen en el nivel inicial. Al identificar que los estudiantes suelen emplear la misma estrategia compensatoria en el estadio inicial y final del período de formación, podemos afirmar, por un lado, que el estudiante percibe la proximidad lingüística como un elemento productivo que le permite comunicarse en L2 y que por eso sigue recurriendo a su L1, y, por otro, que el estudiante no consigue establecer los límites entre su L1 y la L2 debido a la similitud entre ambas, porque el hecho de que se hayan identificado en el nivel avanzado errores con origen en la L1, es un indicio de un posible estancamiento en el desarrollo de su interlengua, tal y como ha señalado Calvo Capilla (2010).

De hecho, la gran proximidad existente entre ambas lenguas y el hecho de que una gran parte de los estudiantes perciba la proximidad lingüística como un elemento facilitador y que ésta aumente a medida que progresan en el aprendizaje de la L2, como se verá más adelante en el apartado 11.7. (La percepción de la proximidad lingüística), son factores que influyen en la persistencia de los errores de origen interlingüístico. Asimismo, a partir de este resultado podemos inferir que los estudiantes siguen aplicando esta misma estrategia a medida que progresa en el aprendizaje del español. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>espetáculo</i>	espetáculo*	espectáculo
<i>assado</i>	assado*	asado
<i>(eu) cresci, crescer</i>	cresci*	(yo) crecí, crecer
<i>cozida</i>	cozida*	cocida
<i>contatos</i>	contatos*	contactos
<i>fria</i>	fria*	fría

En cuanto a los **errores intralingüísticos**, advertimos que su ocurrencia no es uniforme a lo largo de los tres niveles, ya que son los más abundantes en el nivel intermedio. Además, se observa disminución significativa en la ocurrencia de estos errores únicamente entre los niveles intermedio y avanzado. Según Durão (2004), la idea de que la L2 tiene que ser diferente de su L1 lleva a los estudiantes brasileños a alejarse lo más posible del portugués, adaptando gráficamente las palabras y también siendo creativos en el intento de comunicar en la lengua meta. Esta actitud lleva muchas veces a la aparición de palabras que no tienen

correspondientes ni en la L1 ni en la L2. La aplicación de esta estrategia también explicaría la no utilización de acentos en palabras que se escriben y se acentúan de la misma manera en ambas lenguas. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>país</i>	país*	país
<i>próximo</i>	proximo*	próximo

En lo que se refiere a la persistencia de los errores intralingüísticos, hemos visto que aunque se haya observado disminución significativa entre el nivel intermedio (M=4,74) y el avanzado (M=3,29), la media de errores por texto en el nivel avanzado sigue elevada, por lo que se podría decir que hay una tendencia a la permanencia de estos errores en la interlengua de los participantes con un nivel de competencia alto durante el período de formación.

Hemos averiguado también que los **errores ambiguos** son muy poco frecuentes (5,18%) en comparación con los interlingüísticos e intralingüísticos. Además, estos errores disminuyen significativamente entre los niveles intermedio (M=0,58) y avanzado (M=0,33). Un ejemplo de ello son las palabras en las que se puede identificar en el error, características propias de ambas lenguas. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>ensino</i>	ensíño*	enseñanza

Para terminar, queremos precisar que justamente por tener características tanto de la L1 como de la L2 y por tratarse de lenguas muy próximas, a veces es muy difícil identificar el origen del error, definirlo y explicarlo. No es el caso de los errores de préstamo interlingüístico, en los que la causa del error es más explícita, dado que la forma de la palabra producida es idéntica a la de su L1, pero en otros casos como en los que el significante es idéntico en ambas lenguas o cuando la palabra con error tiene características de ambas lenguas, la semejanza muchas veces impide determinar la causa específica del error. Es por ello que estamos de acuerdo con Durão (2004) cuando afirma que hay que seguir investigando sobre el aprendizaje de español por brasileños para profundizar en las estrategias de producción que utilizan los estudiantes de lenguas próximas y averiguar de qué manera variables como el número de horas de instrucción, el tipo de instrucción, el

*input*, la motivación, la percepción de la proximidad lingüística, entre otros, inciden en los resultados.

## 11.6. Los errores léxicos según el tipo y el origen

**Pregunta 6** ¿Cambia el tipo del error léxico (formal, semántico o mixto) según el origen (interlingüístico, intralingüístico o ambiguo) en la producción escrita de los estudiantes brasileños durante el período de formación?

**Resultados:** Los errores interlingüísticos formales disminuyen significativamente entre los niveles inicial e intermedio, pero esta disminución no es estadísticamente significativa entre los niveles intermedio y avanzado. En el caso de los errores interlingüísticos semánticos e intralingüísticos semánticos, no se observa diferencia significativa entre ninguno de los tres niveles. La presencia del error de tipo semántico es constante a lo largo del período de formación, aunque es menor que la de los errores formales. En lo que se refiere a los errores interlingüísticos mixtos e intralingüísticos formales, sólo se observa mejora cuando los estudiantes alcanzan el nivel avanzado. En el caso de los errores intralingüísticos mixtos, la mejora que se observa cuando los estudiantes alcanzan el nivel avanzado no es estadísticamente significativa con respecto al nivel inicial. El análisis de los errores de origen ambiguo, sólo revela mejora en el caso de los errores ambiguos mixtos a partir del nivel avanzado. Además, se constata que este error es muy escaso y tanto en el caso de los errores ambiguos formales como semánticos, no se observa diferencia estadísticamente significativa entre los niveles.

En este apartado discutimos los resultados del análisis de los errores léxicos según el tipo y el origen analizados conjuntamente (Palapanidi, 2010) con el objetivo de profundizar en el conocimiento de las características específicas de los errores presentes en la interlengua de los participantes de este estudio. Antes de empezar a discutir los resultados encontrados en cada uno de los niveles, comentamos las características generales de los datos recogidos en esta investigación, focalizando en los errores más abundantes.

El análisis global de los resultados de los errores según el tipo y el origen indica que los errores más frecuentes en la interlengua de los participantes son los **errores formales** tanto **interlingüísticos** (35,3%) como **intralingüísticos** (35,7%). Asimismo, se puede observar que la proporción de estos dos conjuntos de errores es bastante equilibrada. A éstos les siguen los **errores interlingüísticos semánticos** (7,6%), los **errores interlingüísticos mixtos** (6,9%) y

los **errores intralingüísticos semánticos** (6,2%) respectivamente, aunque en cantidades considerablemente inferiores.

Al analizar los resultados de los errores según el origen y tipo en cada uno de los tres niveles, se observan comportamientos diferentes. En el nivel inicial, los errores interlingüísticos formales presentan la media más alta de errores por texto (M=4,4), y que los errores formales, pero de origen intralingüístico, en cambio, son los que tienen una mayor presencia en los niveles intermedio (M=3,7) y avanzado (M=2,6). Estos resultados revelan características específicas de la interlengua de los estudiantes brasileños de ELE, un contexto de aprendizaje de lenguas afines, que difieren, por ejemplo, de las características de la interlengua de estudiantes de ELE cuya L1 no es tipológicamente similar al español (L2).

En un estudio sobre el aprendizaje de ELE por hablantes griegos, Palapanidi (2010) encontró resultados diferentes de los que hallamos en la presente investigación. En su trabajo, la investigadora observó que la mayoría de los errores léxicos identificados era de errores de origen intralingüístico y de tipo semántico, seguido de los errores intralingüísticos de tipo formal. Creemos que esa discrepancia de resultados se debe a que cuanto mayor la proximidad entre las lenguas implicadas, mayor el número de errores ortográficos con origen en la L1 (Ringbom, 2006), lo que pone de manifiesto que el aprendizaje de una L2 en un contexto de lenguas próximas, como son el portugués (L1) y el español (L2), tiene características propias que lo distancian de los demás.

Tras ofrecer una visión general del comportamiento de los errores léxicos según el tipo y el origen, pasamos a discutir los resultados del análisis de estos errores en cada uno de los tres niveles de aprendizaje. Se ha visto que los errores formales disminuyen significativamente a lo largo del período de formación, así como los errores de origen interlingüístico. Con todo, el análisis de los errores léxicos **interlingüísticos formales** en conjunto, revela una disminución estadísticamente significativa de estos errores solamente entre los niveles inicial e intermedio, lo que indica que cuando el estudiante alcanza el nivel intermedio es capaz de percibir que algunas características formales del portugués (L1) no corresponden a las del español (L2), por lo que ya no se encuentran tantos errores de préstamo como en el nivel inicial. Sin embargo, esta mejora significativa observada en esta fase del aprendizaje

queda estancada, dado que entre el nivel intermedio y avanzado ya no se da una disminución significativa de los errores, lo que indica que los estudiantes siguen recurriendo a su L1 y aplicando la transferencia como estrategia compensatoria en la producción escrita, pero ya no a través de los préstamos como en el nivel inicial, pero a través de transferencias parciales.

A continuación, presentamos ejemplos de errores identificados en la interlengua de los participantes en los que ítems presentan características de la L1. En los tres primeros ejemplos (ej.: *assado\**, *fictício\**, *espetáculo\**), los vocablos empleados en la interlengua coinciden con las formas propias de la L1, mientras que en los demás (*extrangero\**, *faculdad\**, *deverían\**, *trabajam\**), se identifica un solapamiento de ambas lenguas, no obstante el error se debe a las características propias de la L1:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>assado</i>	<i>assado*</i>	asado
<i>fictício</i>	<i>fictício*</i>	ficticio
<i>espetáculo</i>	<i>espetáculo*</i>	espectáculo
<i>estrangero</i>	<i>extrangero*</i>	extranjero
<i>faculdade</i>	<i>faculdad*</i>	facultad
<i>(eles) deveriam, dever</i>	<i>deverían*</i>	(ellos) deberían, deber
<i>(eles) trabalham, trabalhar</i>	<i>trabajam*</i>	(ellos) trabajan, trabajar

Hemos constatado también que los errores interlingüísticos de tipo formal son mucho más frecuentes que los interlingüísticos de tipo semántico o mixto, como se ha visto anteriormente. Dado el elevado número de errores formales que tienen su origen en el portugués (L1) queda patente el uso de la transferencia como estrategia empleada en el intento de escribir en español (L2). Aún así, se ha observado que la transferencia incide de manera distinta a lo largo del proceso de formación, una vez que en el nivel inicial se detectaron más préstamos, mientras que en los niveles intermedio y avanzado se detectaron más ejemplos de transferencias parciales, en los que identifican las formas híbridas, es decir, ítems léxicos que se aproximan más a la L2, pero cuya forma todavía tiene rasgos propios de la L1 que causan el error.

Estos resultados ponen de manifiesto que es necesario focalizar la atención en la forma de las palabras léxicas en la enseñanza de ELE a lusohablantes. La promoción de una enseñanza

contrastiva en contextos de lenguas próximas (ej.: Jordan, 1991; Greco, 2006) no es algo nuevo. Asimismo, creemos que una enseñanza contrastiva asociada a prácticas que fomenten ejercicios de traducción como proponen Calvo Capilla y Ridd (2009), en los que se trabajen la lengua en contexto, puede incidir positivamente en la superación de los errores formales interlingüísticos. Esta práctica pedagógica, asociada a una enseñanza comunicativa, ayudaría en el desarrollo de una conciencia de las características específicas y propias de cada una de las lenguas implicadas, el portugués (L1) y el español (L2), ofreciendo al alumno la oportunidad de progresar en el aprendizaje de la lengua meta de manera eficaz. Asimismo, aunque estos errores no comprometan el significado del mensaje la mayoría de veces, es importante enfatizar que un acento gráfico mal puesto o una letra de más o de menos en una palabra pueden indicar que el estudiante sigue escribiendo en su propia L1, debido al gran parecido de ambas lenguas, y que por lo tanto no está avanzando en el aprendizaje de la L2.

En cuanto a los errores **interlingüísticos semánticos**, la frecuencia de esos errores es muy baja (7,6%) en comparación con los interlingüísticos formales como se ha visto más arriba. No obstante, la presencia de los pocos que se dan es constante, dado que no se observa diferencia estadísticamente significativa entre los niveles. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>oferecer aulas</i>	ofrecer <b>aulas</b> *	ofrecer clases

En el ejemplo vemos el uso del falso cognado ‘aula’ que en español se suele utilizar para referirse al espacio físico en un centro de enseñanza en el que se imparten clases (Marzano, 2006). Sin embargo, el contexto revela que el sustantivo ‘aula’ se ha empleado en lugar de la palabra española ‘clase’, dado que el texto se refería a la enseñanza que se imparte sobre una determinada materia. El par de sustantivos españoles ‘aula’ y ‘clase’ suele confundir al estudiante lusohablante porque también existe en portugués, pero con el significado contrario, como ilustramos en la figura 28.

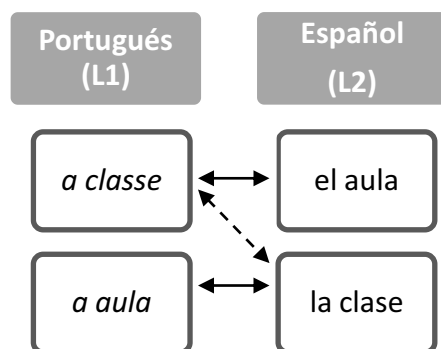


Figura 28. Ejemplo de falsos cognados.

De todos modos, es importante resaltar que una de las acepciones de la palabra ‘clase’ en español se refiere también al espacio físico (ej.: *DRAE*, 2001; Sánchez Pérez, 2001).

Aunque Briones (2000) atribuya los errores más comunes en la traducción entre el portugués y el español, y viceversa, al vocabulario heterosemántico, esto no se ha identificado en los datos analizados en esta investigación. A la luz de estos resultados, podríamos inferir que los participantes demuestran tener conocimiento del vocabulario heterosemántico probablemente porque en la enseñanza preparatoria para la selectividad en Brasil se acostumbra a utilizar gramáticas y material didáctico que incluyen listas de falsos amigos. Así, es posible que tanto el conocimiento adquirido a través del aprendizaje de listas de vocabulario heterosemántico (ej.: Marrone, 2005: 58-59; Marzano, 2006) como la proximidad lingüística y el léxico común que comparten ambas lenguas (Israel, 2001) expliquen la baja ocurrencia de los errores semánticos con origen en la L1. Asimismo, la baja frecuencia de estos errores nos impide hablar de persistencia en la interlengua de este grupo de estudiantes, porque probablemente se tratan de errores individuales.

En el caso de los errores **interlingüísticos mixtos**, se ha visto que además de ser poco frecuentes (6,9%), se nota una disminución estadísticamente significativa en su frecuencia de aparición entre los niveles intermedio ( $M=0,69$ ) y avanzado ( $M=0,42$ ). A modo de ejemplo, comentamos el empleo de la palabra ‘vacas’ en lugar del término ‘plazas’:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>vagas</i>	vacas*	plazas

Este error revela la aplicación de una estrategia en la que el estudiante recurre a la palabra de la L1 y la transforma de manera que se aproxime a la idea que tiene de la L2, lo que



resulta en una palabra que existe tanto en la L1 como en la L2, cuyo significado no tiene relación con el contexto en el que está insertada, dado que el texto hablaba de la necesidad de que hubiera más plazas en los hospitales, por lo que se puede desconsiderar que se estuviera refiriendo a 'vacas', animal mamífero rumiante, hembra del toro. Tal y como se ha visto en el caso de los errores interlingüísticos semánticos, se tratan de errores individuales, es decir, idiosincráticos, que no presentan tendencia a persistir en la interlengua de este grupo de participantes.

Los errores de origen **intralingüístico de tipo formal** comprenden el 35,7% de los errores en el *corpus*, como se ha visto más arriba. Además, al analizar las medias de errores por texto en cada nivel, constatamos que son los errores más numerosos en los niveles intermedio (M=3,7) y en el avanzado (M=2,6), aunque se haya verificado disminución estadísticamente significativa entre ellos. A partir de este resultado, vemos que, pese a que se trate del aprendizaje de ELE en un contexto de lenguas afines, en el que las lenguas implicadas comparten una gran parte del léxico, un mayor número de horas de instrucción es necesario para que se identifique una disminución significativa en los errores intralingüísticos formales, que en este caso ocurre sólo cuando los estudiantes alcanzan el nivel avanzado. Un mayor número de horas de instrucción representa más oportunidades para que los estudiantes experimenten sus hipótesis sobre la L2 y las ajusten de manera adecuada. Muchas veces, tanto alumnos como profesores, asumen que no es necesario practicar determinado vocabulario o estructura, debido a la gran semejanza existente entre el portugués (L1) y el español (L2) que les permite a los estudiantes reconocer y entender dicho vocabulario. Sin embargo, para que lleguen a producir L2 correctamente son necesarias más oportunidades de práctica y consecuentemente más horas de contacto con la lengua meta.

Pese a que se haya observado mejora cuando los alumnos alcanzan el nivel avanzado, consideramos que la media de errores por texto sigue siendo elevada (M=2,6), por lo que consecuentemente los errores intralingüísticos formales manifiestan una tendencia a persistir en la interlengua de los estudiantes durante el período de formación. Además, estos errores son más bien propios de un nivel inicial, como son por ejemplo los errores de acentuación gráfica, de confusión ortográfica y del uso de la estrategia de la diptongación,

que a nuestro parecer ya no deberían figurar en la producción lingüística de un nivel de competencia alto. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>países</i>	países*	países
<i>torcida</i>	afición*	afición
<i>pensado</i>	piensado*	pensado

En cuanto a los errores **intralingüísticos semánticos**, tal y como se ha visto anteriormente en el análisis de los errores de tipo semántico, son poco frecuentes y no se observa ningún cambio estadísticamente significativo entre los tres niveles. Creemos que la baja frecuencia de estos errores reside en la proximidad lingüística existente entre las lenguas implicadas que permite al estudiante lusohablante de español acceder a un abanico de posibilidades a la hora de expresarse en la lengua meta a partir del conocimiento que tiene de su L1. Con todo, el hecho de que este error no haya desaparecido del todo cuando los alumnos alcanzan el nivel avanzado (M=0,47) refleja que siguen experimentando sobre el aprendizaje de reglas y aplicando estrategias como la sobregeneralización o la aplicación incompleta de reglas de la L2. Ejemplo:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>brincar</i>	Tuve la oportunidad de <b>bromear*</b> cuando niño.	jugar

El error que seleccionamos viene dado por el carácter polisémico de la palabra '*brincar*' en portugués (L1) que puede referirse en español tanto a 'bromear' como también a 'jugar', cuando ésta se refiere a "entretenerse o divertirse alguien haciendo algo" (Sánchez Pérez, 2007). El empleo de la palabra 'bromear' en el ejemplo arriba, revela un conocimiento incompleto del vocablo español.

En el caso de los errores **intralingüísticos mixtos**, la mejora que se observa entre los niveles intermedio (M=0,30) y avanzado (M=0,22) no es estadísticamente significativa con respecto al nivel inicial (M=0,33). Con todo, debido a que la ocurrencia de estos errores haya sido muy escasa, podemos afirmar que se tratan más bien de errores idiosincráticos. Eso implica que son más difíciles de identificar durante el período de formación y por lo tanto, más difíciles de tratar. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>próxima</i>	la prójima* vez	próxima
<i>saudade</i>	extrañanza*	nostalgia

Vemos en el primero ejemplo una vez más un caso de polisemia, en el que la palabra portuguesa ‘*próximo*’ que corresponde en español tanto al adjetivo ‘próximo’ (“cercano” *DRAE*, 2001) como al sustantivo ‘prójimo’ (“hombre respecto de otro” Marzano, 2006), ocasiona el error formal que implica consecuentemente un cambio de significado. La confusión de sentidos ocasionada por el error formal es acentuada por el empleo de la forma femenina ‘prójima’ que tiene, además, connotación peyorativa, y que el *DRAE* (2001) define como “mujer de poca estimación pública o de dudosa conducta”. En el segundo ítem, encontramos en la palabra ‘extrañanza\*’ un ejemplo de acuñación en el que ambas lenguas influyen en la creación de un término nuevo. En un primer momento podríamos inferir que se trata de un error de forma cuyo significante correspondiente sería la palabra española ‘extrañeza’. Con todo, el contexto indica que se trata nuevamente de la palabra portuguesa ‘*saudade*’, como se ha visto anteriormente en la discusión de los errores de tipo semántico. La idea de ‘*sentir saudade*’ se suele traducir al español como ‘extrañar’ o ‘añorar’ y el sustantivo ‘*saudade*’ se traduce como ‘añoranza’ o ‘nostalgia’. Así, pensamos que una confusión entre las palabras españolas ‘extrañar’ y ‘añoranza’ es probablemente la causa de este error.

Finalmente, en el análisis de los errores de origen ambiguo tanto en el caso de los errores **ambiguos formales** como de los **ambiguos semánticos**, no observamos diferencia estadísticamente significativa entre los niveles. El análisis de los resultados sólo revela mejora en el caso de los errores **ambiguos mixtos** entre los niveles intermedio y avanzado. Además, constatamos que los tres tipos de error de origen ambiguo son muy escasos a lo largo de todo el período de formación y que las medias en el nivel avanzado son de  $M=0,18$ ;  $M=0,4$  y  $M=0,11$ , respectivamente, por lo que podemos afirmar que se tratan de errores idiosincráticos. Ejemplos:

Portugués (L1)	Error (interlengua)	Español (L2)
<i>lá</i>	iba para alá*	allá
<i>tornar-se</i>	venga a volverme* un profesor	me convierta en, convertirse
<i>qualidade</i>	profesionales de cualidad*	calidad

En el primer ejemplo, tenemos un error ortográfico de omisión de la letra 'l', por lo tanto de tipo formal, y origen ambiguo, porque, dada la similitud entre las formas portuguesa y española es difícil identificar si el ítem con error se debe a la adquisición incompleta de la forma de la lengua meta o si se trata de una adaptación de la forma de la L1. En el segundo, vemos el uso del verbo de cambio 'volverse' en un contexto inapropiado, ya que la idea que transmite este verbo suele implicar no voluntariedad en el cambio, no se acostumbra a usar con profesiones y tiene además un matiz negativo (Milani, 1999), por lo que tenemos un error de tipo semántico. Este error se ha originado tanto en una correspondencia equivocada entre el ítem de la L1 y la traducción a la L2, como en el conocimiento incompleto de la L2 y de los matices implicados en el uso del referido verbo. Esto puede deberse a que los diccionarios sólo ofrecen explicaciones limitadas que no incluyen las características citadas anteriormente que explican por qué el ejemplo seleccionado es un error (volverse: "transformarse alguien o algo en una cosa distinta de lo que era o experimentar alguien un cierto cambio" Sánchez Pérez, 2001). En el tercer y último ejemplo, examinamos un error mixto (formal y semántico) y ambiguo (inter e intralingüístico), en el que tenemos la palabra polisémica portuguesa '*qualidade*' que puede referirse en español tanto al vocablo 'cualidad' como a 'calidad' (Marzano, 2006). Así, vemos a partir del ejemplo, como un ítem bastante similar en ambas lenguas comporta una serie de características que dificultan establecer si se trata simplemente de un error semántico que tiene origen en la forma de la palabra de la L1, o si es un error intralingüístico mixto porque la forma utilizada corresponde a un ítem de la L2, pero no el adecuado al contexto en el que está siendo empleado. De todos modos, creemos que se trata de un error semántico y formal, por lo tanto mixto, ocasionado por la polisemia del ítem de la L1 que tiene dos referentes en la L2, por lo tanto ambiguo.

Para finalizar, del análisis de los errores de tipo y origen en conjunto, destacamos por un lado la alta frecuencia de los errores formales tanto interlingüísticos como intralingüísticos, y por otro, pese a ser muy baja, la ocurrencia constante de los errores semánticos interlingüísticos e intralingüísticos durante el período de formación. Como se ha dicho anteriormente, creemos que la proximidad lingüística, bajo la perspectiva del análisis de los errores, ejerce un papel fundamental en estos resultados porque por un lado incide de manera negativa en el aspecto formal, causando confusión a lo largo de todo el proceso de

formación, y por otro, de manera positiva en el aspecto semántico, ya que ambas lenguas comparten una amplia parte del léxico, lo que explicaría la baja frecuencia de este tipo de error. Además, los errores identificados no suelen interrumpir la comunicación en la mayoría de los casos, lo que puede hacer que persistan en la interlengua de los estudiantes, dado que, según el criterio de prioridad en la corrección de Olsson (1973), estos estarían en el final de la lista de prioridades de corrección del profesor en un contexto de enseñanza con enfoque comunicativo.

### 11.7. La percepción de la proximidad lingüística

Las lenguas implicadas en el proceso de aprendizaje analizado en esta investigación, el portugués (L1) y el español (L2), son lenguas latinas que comparten una gran parte del léxico (Israel, 2001). Esta proximidad lingüística tiene implicaciones en el proceso de enseñanza-aprendizaje como, por ejemplo, el hecho de que la inteligibilidad permite que la comunicación entre el profesor y los alumnos se realice en la L2 desde el primer día de clase (Camorlinga, 1997; Espiga, 1999; Natel, 2002; Fialho, 2005). Otra implicación derivada de la semejanza tipológica es que ésta favorece la transferencia léxica (Cenoz, 2001) en comparación con contextos de aprendizaje en los que la L2 no tiene relación de proximidad con la L1 de los estudiantes. Aunque la transferencia como estrategia tiende a disminuir a medida que aumenta el número de horas de instrucción (Celaya y Torras, 2001; Woodall, 2002; Navés *et al.*, 2005; Celaya, 2007), la percepción que tienen los estudiantes de la distancia tipológica entre la L1 y la L2 (psicotipología) (Kellerman, 1978, 1986) es un factor clave que influye en los procesos que determinan la transferencia, así como la percepción de la especificidad de un determinado ítem en la L1 (prototipicalidad) y el nivel de competencia que tiene el estudiante en la L2. Así, vemos que no sólo la proximidad lingüística en sí, sino la percepción que el estudiante tiene de la proximidad, es uno de los factores que influye en las estrategias de producción y consecuentemente en el origen de los errores.

La proximidad tipológica en el contexto de aprendizaje de lenguas afines implica una dicotomía porque se la puede considerar tanto de forma positiva, una vez que facilita y agiliza el aprendizaje de la L2, como negativa, porque la semejanza también induce a los

errores (Calvi, 1998; Andrade Neta, 2000). Nuestra experiencia como alumna y profesora de ELE a brasileños ha motivado a explorar más a fondo qué percepción de la proximidad lingüística tienen los estudiantes en este contexto y averiguar si ésta cambia a lo largo del período de formación. Así, se podría analizar cómo interactúan la percepción de la proximidad y el nivel de conocimiento de los estudiantes en la producción de errores léxicos en el ámbito de la enseñanza de ELE a brasileños.

A partir del análisis de las respuestas obtenidas en el cuestionario sobre la percepción de la proximidad lingüística (anexo 4), hemos averiguado que, de manera general, la percepción de la proximidad entre el español (L2) y el portugués (L1) como un elemento facilitador de su aprendizaje está más extendida que la percepción de la semejanza como un factor que añade dificultad al aprendizaje. Con todo, el análisis de los resultados en cada uno de los tres niveles revela que esta percepción aumenta a medida que los estudiantes mejoran su competencia en la L2. Así, vemos que en el nivel inicial el 44% de los participantes ve la proximidad tipológica como una característica que les ayuda en el aprendizaje de la L2, mientras que en el nivel intermedio este porcentaje sube al 55% y en el avanzado pasa al 64%.

Sin embargo, se ha visto que tanto los estudiantes que piensan que aprender español es fácil como los que piensan que es difícil citan la semejanza entre las lenguas portuguesa (L1) y española (L2) como el principal motivo de su apreciación. Así, aunque una mayor parte de los estudiantes percibe la proximidad lingüística como un elemento positivo que les ayuda en su proceso de aprendizaje, los demás creen que la causa de sus problemas está justamente en la semejanza existente entre la L1 y la L2.

Además, se ha visto que los estudiantes perciben que las destrezas productivas (escribir y hablar) les suponen más dificultad que las receptivas (leer y escuchar). Esto se debe posiblemente a que el dominio del vocabulario pasivo, aquél que el estudiante es capaz de reconocer, es mayor que el del activo, aquél que el estudiante es capaz de producir (Laufer, 1998, 2003). Asimismo, en el caso específico de la expresión escrita, los estudiantes consideran que sus capacidades en esta destreza mejoran a medida que ellos avanzan en el proceso de aprendizaje de la L2. Los resultados discutidos en los apartados anteriores están en línea con esta percepción de los estudiantes ya que se ha visto una mejora de la

diversidad léxica y la disminución significativa de errores léxicos, ambas cuando los estudiantes alcanzan el nivel avanzado.

En lo que se refiere a la tarea propuesta, el 61% de los participantes afirmó haber tenido dificultad para escribir las redacciones y citaron el vocabulario y la ortografía como causas principales de su dificultad. Este resultado está acorde con lo que hemos visto en el apartado 11.4. (Los errores léxicos según el tipo), dado que los errores formales son los más frecuentes en la muestra, lo que revela que los estudiantes son conscientes de sus dificultades.

El objetivo del aprendizaje de una L2 hoy en día, y que compartimos, suele ser desarrollar la capacidad de comunicación del estudiante en la lengua meta. Es por ello que los docentes de ELE en Brasil se enfrentan frecuentemente al pensamiento que comparten muchos alumnos que consideran que no hace falta estudiar la lengua española para ser capaces de comunicarse. Señalamos, con todo, que esta idea generalizada de que es posible comunicarse sin conocer la lengua española, no siempre es compartida por los interlocutores hispanohablantes. Además, cuando se establece comunicación, muchas veces falla la corrección. Creemos que tanto la proximidad lingüística, la percepción de la proximidad como elemento facilitador del aprendizaje y de la comunicación, como también el carácter generalmente extrovertido del hablante brasileño, que suele lanzarse y arriesgarse en situaciones comunicativas en lengua española aunque no la conozca muy bien son factores que contribuyen para que la corrección quede en un segundo plano. Con relación a este último factor, Durão (2004) observa, a partir de los resultados de su investigación sobre los errores que producen un grupo de lusohablantes brasileños estudiantes de español y uno de hispanohablantes españoles estudiantes de portugués, que pese a compartir la proximidad tipológica, los brasileños y los españoles forman, de manera general, dos grupos que tienen actitudes distintas frente al aprendizaje de la L2 y la autora lo atribuye a las características socioculturales propias de cada uno.

En nuestro análisis evidenciamos una baja frecuencia de errores semánticos, probablemente porque los temas propuestos tenían el propósito de estimular la producción escrita y por ello se referían a temas de ámbitos cercanos al alumno. Este resultado coincide con la idea extendida de que hay comunicación efectiva entre hablantes del portugués (L1) y hablantes

de español (L2) en determinadas situaciones de intercambio lingüístico. Situación semejante a la que ocurre con otro par de lenguas afines, el español y el italiano, como explica Calvi (2004). Con todo, debemos tener presente que la comunicación efectiva depende de muchos factores como son los interlocutores y su actitud, el contexto, la utilización de recursos comunicativos, gestos, entre otros.

Sin embargo, el elevado número de errores formales identificados en el *corpus* afecta a la corrección y muestra que el estudiante y futuro profesor de ELE no tiene un dominio de las características propias de la L2. Por un lado, el elevado número de errores de origen interlingüístico especialmente en el nivel inicial, pero también presentes en los niveles intermedio y avanzado, revelan que los estudiantes recurren a su L1 como estrategia compensatoria a lo largo de todo el período de formación, por lo que intuimos que existe una relación entre su percepción de la proximidad como elemento facilitador y la persistencia de los errores interlingüísticos en un nivel de competencia alto.



## PARTE V: CONCLUSIONES Y CONSIDERACIONES FINALES

### CONCLUSIONES

La quinta y última parte de esta tesis está formada por las conclusiones generales a las que hemos llegado a partir del análisis y la discusión de los resultados. Con el objetivo de facilitar la lectura, las organizamos en diferentes apartados. Primeramente, presentamos las conclusiones que hemos obtenido tras responder a las preguntas de investigación. A continuación, proponemos algunas implicaciones didácticas como contribución al campo de la didáctica de la enseñanza del español como lengua extranjera (ELE) en contextos de lenguas afines a partir de la reflexión que ha generado esta tesis. Por último, señalamos algunas limitaciones que pueden atañer este estudio e indicamos algunas líneas para investigaciones futuras.

El objetivo general de esta tesis es conocer las características de la interlengua en un contexto de aprendizaje de lenguas próximas o afines, en este caso, de ELE por hablantes de portugués (L1). Para ello, examinamos los errores léxicos identificados en la producción escrita de estudiantes universitarios brasileños en formación como profesores de ELE.

Para profundizar en el análisis de los errores léxicos, los clasificamos según el tipo, es decir, averiguamos si se trataban de errores que afectaban a la forma o al significado de las palabras. Examinamos también el origen del error, es decir, si los errores léxicos se debían a la influencia del portugués (origen interlingüístico), o si se debían al proceso de desarrollo de la interlengua (origen intralingüístico). Además, nos interesaba especialmente saber si se observaban cambios en la frecuencia de esos errores en tres diferentes momentos del período de formación, es decir, si los estudiantes de un nivel de competencia alto presentaban menos errores léxicos en su producción escrita, y si el tipo y el origen de esos errores eran diferentes de los identificados en la producción de los estudiantes de un nivel de competencia bajo. Además, analizamos los errores de tipo y origen en conjunto, para tener una visión más completa del comportamiento de los errores léxicos en este contexto y comprobar si se daban cambios al comparar los resultados de los diferentes niveles de competencia. También exploramos si había una tendencia a la fosilización/estabilización de

algunos errores léxicos en la interlengua de los estudiantes dada la proximidad de las dos lenguas. Así, por un lado, hemos querido comprobar si algunos de esos errores persistían en su interlengua y, por otro lado, si algunas de las estrategias utilizadas que generaron esos errores podrían persistir en la producción de su interlengua a lo largo del período de formación. Finalmente, exploramos la percepción que tenían los estudiantes de la proximidad lingüística entre el español (L2) y el portugués (L1) para averiguar si la percibían como un elemento facilitador o no del aprendizaje de la lengua meta.

Para la investigación empírica, optamos por la aplicación de una metodología cuantitativa y un diseño pseudolongitudinal. Para ello diseñamos varias pruebas para recoger datos lingüísticos escritos que aplicamos a una muestra de un total de 142 alumnos de lengua española del Departamento de Letras de la *Universidade Federal de Pernambuco* (UFPE), Brasil. Tras la recogida de datos, las 365 producciones escritas obtenidas se agruparon en tres bloques de análisis, según el nivel de competencia y el número de horas de asignaturas de lengua española que habían cursado los participantes. Se identificaron un total de 4.407 errores entre las palabras léxicas y funcionales. Posteriormente se clasificaron los 3.302 errores léxicos según la clase de palabra léxica a la que afectaban (sustantivos, adjetivos, verbos, y adverbios), tipo de error léxico (formales, semánticos y mixtos) y origen del error léxico (interlingüísticos, intralingüísticos y ambiguos). El análisis de los datos se realizó principalmente a partir de los resultados de pruebas estadísticas de tipo no paramétrico, seguido de un análisis cualitativo complementario que procuraba examinar más a fondo las características de los errores.

A continuación desglosamos las conclusiones a las que hemos llegado a partir del análisis y la discusión de los resultados en los siguientes apartados:

#### 1) La fluidez y la diversidad léxica

La fluidez y la diversidad léxica aumentan a medida que aumenta el número de horas de instrucción, pero no hay un paralelismo entre ambas. La fluidez aumenta en el nivel intermedio y se estanca a partir de este momento, mientras que la diversidad léxica muestra un aumento en las producciones del nivel avanzado, es decir, los alumnos del nivel avanzado no escriben más que los del nivel intermedio, pero lo hacen con un

vocabulario más variado y correcto, dado que la medida aplicada para el análisis de la diversidad léxica excluía los errores.

El alto grado de fluidez que se ha observado al principio del período de formación lo atribuimos a la semejanza entre las lenguas portuguesa (L1) y española (L2), que permite al estudiante aproximarse a la L2 a partir del conocimiento previo de su L1, el cual le ofrece las herramientas necesarias para expresarse y comunicarse, pero al precio de un gran número de errores. En cambio, el aumento de la riqueza léxica detectada en el nivel avanzado se debe a que, tras un mayor número de horas de instrucción, los participantes demuestran haber incorporado y ser capaces de utilizar una mayor variedad de palabras léxicas correctamente.

## 2) Los errores léxicos y funcionales

El número de errores léxicos es mayor que el número de errores funcionales a lo largo de todo el período de formación. Además, ambos errores disminuyen significativamente cuando los estudiantes alcanzan el nivel avanzado, pero sin que se lleguen a erradicar. Pese a que se haya identificado una tendencia a la persistencia del error léxico en las producciones de los estudiantes con un nivel de competencia alto, la calidad de los textos escritos, medida a través de la riqueza léxica y de la corrección, mejora considerablemente cuando los estudiantes alcanzan dicho nivel. Así, vemos que en un contexto de lenguas afines, un mayor número de horas de instrucción y contacto con la L2 influye de manera positiva en la mejora de la calidad de la producción escrita en la lengua meta.

## 3) Los errores según la clase de palabra léxica

La media de palabras léxicas utilizadas por texto aumenta a medida que aumenta el número de horas de instrucción y este comportamiento se repite en cada una de las clases de palabra léxica (sustantivos, adjetivos, verbos y adverbios) cuando se computan por separado. Se han identificado errores en todas las clases de palabra léxica, pero con comportamientos distintos a lo largo del período de formación.

Respecto a los sustantivos, si bien su uso aumenta a lo largo de los tres niveles, la presencia de errores se mantiene, y sólo se observa diferencia significativa cuando comparamos las medias de errores entre los niveles inicial y avanzado. En segundo lugar, en cuanto a los verbos, también aumentan su uso, pero son los que presentan una incidencia de errores más alta tanto en el nivel inicial como intermedio, aunque disminuyen progresivamente a medida que aumenta el número de horas de instrucción y sobretodo en el nivel avanzado. En cuanto a los adjetivos y adverbios, se utilizan menos que las clases anteriores. Los errores en los adjetivos y adverbios son menos frecuentes que los demás y disminuyen cuando los estudiantes alcanzan el nivel avanzado.

#### 4) Los errores léxicos según el tipo: formales, semánticos o mixtos

Los errores formales, es decir, aquéllos que afectan a la representación gráfica de la palabra, son mucho más numerosos que los errores semánticos, que afectan al significado, o mixtos, que se ven alterados tanto en la forma como en el significado. Aunque los errores formales disminuyen significativamente a lo largo del período de formación, siguen siendo los más abundantes en los textos del nivel avanzado.

Los errores formales afectan a todos los tipos de palabras léxicas y presentan una gran variedad. Esos errores, aunque disminuyen significativamente su presencia en el nivel avanzado, no se erradican, e inciden en la ortografía, como la omisión y adición de acentos; adición, omisión o sustitución de letras que conforman la palabra; confusión de grafías motivadas fonológicamente (fonema /b/, grafías 'v' y 'b'; fonema /s/, grafías 'ss' y 's'; fonema /k/, grafías 'qu' y 'c'; fonema /x/, grafías 'j' y 'g'), y en el uso frecuente y erróneo de la diptongación.

De los errores semánticos y mixtos, aunque mucho menos frecuentes, pero siempre presentes a lo largo del proceso de formación, destacan los falsos cognados, los debidos a la polisemia del ítem de la L1 y acuñaciones en los que se mezclan características de ambas lenguas.

## 5) Los errores léxicos según el origen: interlingüísticos, intralingüísticos y ambiguos

Se ha verificado la presencia en los tres niveles tanto de errores interlingüísticos, aquéllos que tienen su origen en la L1, como intralingüísticos, aquéllos que se deben al propio proceso de desarrollo de la interlengua, e incluso de los ambiguos, aquéllos que presentan las dos influencias.

Al considerar los datos globalmente, vemos que los errores interlingüísticos superan a los intralingüísticos. Se constata que en el nivel inicial, tal y como indican investigaciones anteriores, tanto en contextos de lenguas tipológicamente próximas como distantes, la frecuencia de errores interlingüísticos es más abundante. En el nivel intermedio, sin embargo, la presencia de estos errores disminuye de manera significativa, lo que indica que los estudiantes recurren menos a su L1 y los errores de origen intralingüístico pasan a ser los más numerosos. Paradójicamente, cuando los alumnos alcanzan el nivel avanzado, pese a que se observa una disminución significativa de esos dos tipos de errores, se verifica una presencia equilibrada de ambos, reflejando que en un nivel avanzado los estudiantes aún tienen dificultad para establecer los límites entre los dos idiomas.

El análisis de los errores ambiguos indica que éstos son muy poco frecuentes y que además disminuyen significativamente cuando los alumnos alcanzan el nivel avanzado, por lo que se tratan probablemente de errores idiosincráticos.

## 6) Los errores léxicos según el tipo y el origen

Los errores formales son los que tienen mayor peso en el *corpus* general y su origen se divide en partes casi iguales entre los que se deben a la influencia de la L1 (error interlingüístico) y los que se deben al proceso de desarrollo de la interlengua (error intralingüístico). Sin embargo, cuando se observa su presencia en cada uno de los niveles la proporción no es la misma. En el estadio inicial los errores formales interlingüísticos superan a los intralingüísticos, al contrario de lo que ocurre en el estadio intermedio en que estos últimos superan a los primeros. En el nivel avanzado,

la presencia de ambos errores se iguala al disminuir los intralingüísticos y mantenerse los que tienen su origen en la L1.

Los errores semánticos, tanto interlingüísticos como intralingüísticos, son poco frecuentes y no se observa cambio entre los tres niveles, pese a que estén presentes a lo largo de todo el período de formación, mientras que los errores mixtos de ambos orígenes presentan una mejora sólo cuando los estudiantes alcanzan el nivel avanzado.

Por último, se ha visto también que los errores de origen ambiguo, tanto formales, como semánticos, son muy escasos y no se observa cambios entre los niveles, excepto en el caso de los errores mixtos que bajan significativamente también cuando los estudiantes alcanzan el nivel avanzado.

#### 7) La persistencia de los errores en el nivel avanzado

A partir del análisis de los errores léxicos identificados en las producciones escritas de los alumnos del nivel avanzado, identificamos ciertas características que pueden indicar una tendencia a la persistencia no sólo de algunos de esos errores en su interlengua, sino también de algunas de las estrategias compensatorias que emplean los alumnos en su intento de aproximarse a la lengua meta.

En primer lugar, vimos que tras el progreso que se observa cuando los participantes alcanzan el nivel intermedio, la fluidez se mantiene invariable entre los niveles intermedio y avanzado.

En segundo lugar, constatamos que los errores léxicos, aunque mejoren significativamente a lo largo del período de formación, siguen siendo muy frecuentes en el nivel avanzado.

En tercer lugar, considerando los errores según la clase de palabra, observamos que en el caso de los sustantivos, no se contemplan cambios en las medias de errores entre los niveles, aunque cabe decir que la producción de sustantivos aumenta a medida que aumenta el número de horas de instrucción. Por lo que respecta a los

verbos, pese a que sean los que presentan más errores cuando se considera el *corpus* globalmente, éstos disminuyen progresivamente a lo largo del periodo de formación, pero no se llegan a erradicar.

En cuarto lugar, vimos que los errores formales, sobretodo en la ortografía, dada su abundante ocurrencia y pese a que se verifique mejora entre los niveles, presenta una tendencia a persistir en la interlengua, posiblemente ocasionada por la gran parte del léxico común y por la gran proximidad en cuanto a muchas de las formas de las palabras del portugués (L1) y del español (L2). Los errores semánticos son poco frecuentes y no presentan mejora entre los niveles.

En quinto y último lugar, cuando se analizan los errores formales según su origen se observa que tanto los que tienen su origen en la L1, sobre todo en las reglas ortográficas de la L1, como los que son resultado de desconocimiento de las reglas ortográficas de la L2, pese a tener un comportamiento diferente, no se erradicar en el nivel avanzado. En el caso de los errores semánticos que son mucho más reducidos, persisten los que tienen su origen en la L1, como es el caso de los errores que se originan en la equivalencia o traducción equivocada entre las dos lenguas debido a polisemia del ítem en la L1.

## 8) Estrategias

El análisis cuantitativo apoyado por la exploración cualitativa y un estudio detallado de los errores formales y semánticos según su origen revela la utilización de diferentes estrategias compensatorias utilizadas por los estudiantes en su intento de producir L2.

Se aprecia que una de las más utilizadas es la transferencia del conocimiento previo de su L1. Con todo, la transferencia se manifiesta de manera diferente según el grado de conocimiento que el estudiante tiene de la L2. En el nivel inicial se detectan más errores de préstamos, es decir, se identifican en la interlengua ítems de la L1 sin ningún tipo de adaptación a la L2 a los que los alumnos recurren para compensar la falta de conocimiento léxico en la lengua meta. A partir del nivel intermedio, y también en el nivel avanzado, se observa un cambio en el empleo de esta estrategia,

ya que no se identifican tantos préstamos como al principio del aprendizaje, y éstos dan lugar a las transferencias parciales, lo que origina formas híbridas, es decir, ítems en los que se identifican características tanto de la L1 como de la L2. Esta mezcla entre los elementos de ambas lenguas que pueden ser ortográficos y/o semánticos es una estrategia creativa que el estudiante aplica en su intento de aproximarse a la L2, algo que hace cuando ya ha alcanzado cierto conocimiento de la lengua meta.

En cuanto a los errores sin influencia de la L1, la diptongación, la hipercorrección y las generalizaciones son estrategias bastante utilizadas en el intento que hace el alumno por acercarse a la L2. La diptongación se refiere a la idea extendida de que en la lengua española hay muchos más diptongos que en la L1, por lo que para producir L2 bastaría con añadir un diptongo a una palabra de la lengua primera, lo que resulta muchas veces en lo que se conoce como ‘portuñol’. El empleo de esta estrategia es mucho más frecuente en los niveles inicial e intermedio, mientras que la hipercorrección y las generalizaciones empiezan a tener más peso a partir del nivel intermedio, cuando la producción escrita de los estudiantes ya no presenta tantos ejemplos de influencia de la L1 como en el nivel inicial.

Tras el cambio que se observa tras finalizar el nivel inicial, el análisis más detallado de los errores identificados a partir del nivel intermedio ha indicado la existencia de una tendencia a la estabilización de las estrategias utilizadas por los alumnos en su intento de producir L2, lo que nos lleva a referirnos a este fenómeno a través del término ‘fossilización de estrategias’, en el sentido que el aprendiz pueda seguir aplicándolas siempre generando el mismo tipo de errores.

## 9) Percepción de la proximidad lingüística

La percepción de los participantes en el estudio sobre el papel que juega la semejanza lingüística entre el portugués (L1) y el español (L2) en el aprendizaje de esta segunda lengua se distribuye en partes iguales entre los que consideran que es un elemento facilitador y aquéllos que la consideran una fuente de dificultad. Con todo, al comparar los datos por nivel, observamos que a medida que aumenta el grado de



conocimiento de la lengua meta, aumenta la percepción de los participantes de que aprender español es fácil.

Apreciamos de ello que la L1 tiene un papel substancial en ambos casos, porque en el primero, los estudiantes sienten que pueden recurrir a su L1 cuando no conocen una palabra específica en L2, lo que puede resultar tanto en lo que se conoce como transferencia positiva como en error o transferencia negativa. En el segundo, los estudiantes son conscientes de que la semejanza les impide progresar porque les cuesta definir claramente qué elementos pertenecen a cada una de las dos lenguas y la conciencia de que la L2 tiene que ser diferente de la L1 puede llevar a la evitación de las formas semejantes a la L1 y a la creación de palabras que no existen en ninguna de las dos lenguas.



## CONSIDERACIONES FINALES

- Implicaciones didácticas

A partir de la reflexión originada por la discusión de los resultados y de las conclusiones de esta investigación que acabamos de exponer, sugerimos algunas propuestas que podrían considerarse en el ámbito de la didáctica de la enseñanza de ELE en contextos de lenguas afines como en el caso de los lusohablantes.

En primer lugar, debido al elevado número de errores formales que tienen su origen tanto en la influencia de la lengua primera (L1) como en el propio proceso de aprendizaje de la lengua segunda (L2), que se ha detectado en los diferentes estadios del período de formación de los participantes, sugerimos que la enseñanza del español para hablantes del portugués incida más en las características formales propias de cada una de las lenguas implicadas. Creemos que una enseñanza contrastiva, al margen de la carga valorativa negativa que conlleva el término, integrada a una enseñanza comunicativa, incidiría de manera positiva en la superación de los errores formales que tienen su origen en la L1 y ayudaría al alumno a utilizar la semejanza lingüística a su favor, una vez que fuera consciente de las características comunes a ambas lenguas, así como de las que son propias de cada una de ellas.

En segundo lugar, pese a que los errores en las traducciones entre lenguas afines como el portugués y el español se deban en su mayoría a los falsos cognados, en esta investigación hemos identificado pocos errores semánticos en comparación con el número de errores formales, probablemente porque existe una conciencia extendida sobre la necesidad de conocer el vocabulario heterosemántico. Aún así, creemos que es necesario el desarrollo de actividades que promuevan el aprendizaje y la ampliación del léxico no sólo a través de situaciones culturales y pragmáticas significativas, sino también de la enseñanza de las colocaciones propias de la L2, para que los estudiantes lusohablantes puedan ser capaces de producir un discurso más fluido y preciso.

En tercer lugar, hemos visto que algunos errores léxicos presentan mejora significativa sólo cuando los alumnos alcanzan el nivel avanzado. Una implicación didáctica derivada de este

hallazgo está directamente relacionada con el diseño de cursos de ELE para lusohablantes y programas de formación, dado que, como se ha visto, los estudiantes, pese a disponer de un conocimiento previo dado por la semejanza tipológica cuando empiezan el aprendizaje, necesitan tiempo y oportunidades para probar sus hipótesis, ajustarlas y asimilar el nuevo conocimiento. El hecho de que los alumnos brasileños avancen rápidamente al principio de su proceso de aprendizaje y de que tengan facilidad para comprender mensajes en español crea la ilusión de que no es necesario dedicarse al estudio de la lengua española. Creemos que este comportamiento asociado a la percepción de facilidad que proporciona la proximidad lingüística son factores que influyen en el ritmo y en el resultado del aprendizaje, por lo que sería necesario promocionar una enseñanza que incida también en la formación de estrategias y actitudes hacia el aprendizaje de la lengua meta, que desarrolle la conciencia de los estudiantes sobre las estrategias de aprendizaje que pueden utilizar para facilitar este proceso, es decir, desarrollar su competencia de aprender a aprender en este contexto de lenguas afines.

La atención a los aspectos formales de la lengua no debe ser obstáculo para que la motivación de los estudiantes baje. Creemos que es necesario promover una enseñanza que se base en la comunicación en situaciones variadas y significativas, que resalte las diferencias existentes entre las dos lenguas implicadas más allá del tópico de los heterosemánticos, y que insista en el hecho de que aunque una determinada palabra exista en ambas lenguas con significado similar, no significa que se usen de la misma manera, ni con la misma frecuencia, ni en las mismas situaciones comunicativas. Consecuentemente, creemos necesario promocionar una enseñanza que focalice tanto en el conocimiento del sistema lingüístico como en los aspectos comunicativos y pragmáticos.

Por todo lo dicho, pensamos que se debe seguir investigando sobre el aprendizaje del léxico en un contexto de lenguas próximas para profundizar en los múltiples factores que influyen en este proceso, como son por ejemplo, la práctica pedagógica, como la enseñanza de estrategias de aprendizaje, los materiales, el *input* y la motivación, y ampliar el conocimiento sobre las características de la interlengua en este contexto de aprendizaje. A partir de ello, se podrá averiguar si las conclusiones de nuestro estudio se pueden extender a otros contextos semejantes, lo que conllevaría a la aplicación de este conocimiento en la elaboración de intervenciones didácticas y materiales específicos con el objetivo de

promover un aprendizaje más consciente y eficaz que revierta en un mejor desempeño lingüístico de los estudiantes que incluya además de las competencias comunicativas un alto grado de corrección.

- Limitaciones

Debemos señalar que esta investigación presenta algunas limitaciones. En primer lugar, dado que utilizamos una metodología pseudolongitudinal, no seguimos a los mismos estudiantes durante su proceso de aprendizaje, y por ello no podemos hablar de evolución, sino de tendencias que se observan en tres momentos diferentes de la formación académica y que ofrecen un panorama indicativo de lo que ocurre a lo largo del período de formación. Sin embargo, al tratarse de un estudio empírico pseudolongitudinal, pudimos conocer los errores léxicos que producen los estudiantes en diferentes momentos de dicho período. Asimismo, el desarrollo y la aplicación de un estudio con estas características nos han permitido acceder a un gran número de participantes, y, consecuentemente, obtener un *corpus* lingüístico importante, utilizable para futuras investigaciones.

En segundo lugar, consideramos que generalizar estos resultados a contextos similares debería hacerse con cautela, puesto que los resultados pueden haber sido influenciados por factores como el instrumento utilizado para la recogida de datos, que proponía el desarrollo de temas previamente seleccionados, o por el hecho de que la tarea se haya realizado en clase, lo que pese a ofrecer una situación de igualdad de condiciones para todos los participantes, puede influir de manera diferente en el resultado, el texto escrito, por tratarse de una situación en la que el tiempo es limitado. Aún así, como se ha explicado en la metodología, la selección de los temas propuestos y las consignas que se dieron a los participantes tenían el propósito de controlar y homogeneizar las condiciones de realización de la actividad, a la vez que procuraba proporcionarles temas interesantes y cercanos que estimularan la producción escrita en una situación semejante a la que estaban acostumbrados a practicar en el contexto de formación universitaria.

Por último, nos gustaría resaltar que la presente investigación se ha realizado sobre una muestra de estudiantes de lengua española del Departamento de Letras de la *Universidad*

*Federal de Pernambuco* (UFPE), Brasil, por lo que la generalización a otros contextos distintos al incluido en nuestros datos debería hacerse con cautela. Esto porque aunque Brasil sea un país monolingüe, las características propias de cada región como variedades sociolingüísticas, regionales y culturales, además de la proximidad con países hispanohablantes, pueden influir de alguna manera en la tipología de errores de los datos.

No obstante, pese a las limitaciones, pensamos que los resultados y las conclusiones presentadas anteriormente contribuyen a un mejor conocimiento de los procesos implicados en la Adquisición de Segundas Lenguas, especialmente en lo que se refiere a los diferentes estadios de la interlengua de estudiantes de ELE en contextos de lenguas afines. Esto se debe principalmente al hecho de que en esta investigación se ha empleado una metodología pseudolongitudinal que nos ha posibilitado conocer un poco más las características propias de los errores léxicos que se identifican en diferentes estadios de la interlengua de estudiantes de ELE en un ámbito de lenguas afines.

- Futuras líneas de investigación

Esta tesis deja abiertas diferentes líneas para investigaciones futuras. En primer lugar, sugerimos realizar investigaciones cualitativas que profundicen en el estudio de las características de los errores léxicos, para explorar más a fondo las estrategias utilizadas por los estudiantes.

En segundo lugar, otra línea de investigación, sería explorar, más profundamente y a través de estudios cualitativos longitudinales, la tendencia que observamos a la persistencia de determinados errores al finalizar el período de formación, para ampliar el conocimiento sobre el fenómeno de la fosilización/estabilización en contextos de lenguas próximas. Es probable que al finalizar sus estudios e iniciar la práctica docente, los estudiantes sigan perfeccionando su conocimiento de la L2. Con todo, sería necesario realizar estudios de seguimiento a través de pruebas diseñadas concretamente para este fin una vez finalizado el período de formación para averiguar si se confirma la tendencia a la fosilización/estabilización de estrategias detectada anteriormente. Ello podría ser utilizado

para los programas de formación continua que deberían contemplar no sólo los aspectos didácticos como también los de perfeccionamiento lingüístico.

Finalmente, sugerimos también el análisis de los errores sintácticos con el objetivo de explorar el comportamiento y la evolución de este tipo de error a lo largo del aprendizaje de ELE. Dado que se trata de un conocimiento más abarcable que el dominio del léxico, una vez que hace parte de un ámbito cerrado, sobre el que hay gramáticas y manuales con reglas concretas a las que los estudiantes pueden recurrir en todo momento, intuimos que se detectarían menos errores que en el ámbito del léxico. Aún así, sería de gran interés identificar si hay estructuras propias de la L2 que los estudiantes evitan y en qué medida recurren a su L1 cuando tienen que escribir en español.

\*\*\*\*\*

Como conclusión final, creemos que este trabajo puede constituir una aportación relevante para el conocimiento de un campo de estudios poco explorado como es el caso del contexto de aprendizaje de lenguas afines, y más específicamente del aprendizaje del español por hablantes de portugués. Si a ello le añadimos que la política lingüística educativa actual en Brasil pone especial énfasis en fomentar la enseñanza del español y, consecuentemente, en la necesidad de formación de profesores de esta lengua, consideramos que las aportaciones de este estudio y las implicaciones didácticas que se puedan derivar pueden contribuir a una mejora de los programas de formación del profesorado.





## BIBLIOGRAFÍA

- AGUDELO, S. P. (2011): *Los métodos de enseñanza en ELE: El método comunicativo revisado*, Memoria de Máster, Université de Montréal.
- AGUSTÍN LLACH, M. P. (2007): "Lexical errors as Writing Quality Predictors", *Studia Linguistica*, 61 (1), 1-19.
- AGUSTÍN LLACH, M. P. (2010): "An Overview of Variables Affecting Lexical Transfer in Writing: A Review Study", *International Journal of Linguistics*, 2 (1), 1-17.
- AGUSTÍN LLACH, M. P. (2011a): *Lexical Errors and Accuracy in Foreign Language Writing*. Bristol: Multilingual Matters.
- AGUSTÍN LLACH, M. P. (2011b): "Una breve panorámica sobre la adquisición del léxico en la lengua extranjera mediante la herramienta del análisis de errores", *Revista Nebrija de Lingüística Aplicada*, 9 (5), 22- 26.
- AINCIBURU, M. C. (2007): *La adquisición del léxico en las lenguas afines: El aprendizaje de cognados, falsos amigos y palabras nuevas por alumnos italianos de Español como Lengua Extranjera*, Tesis Doctoral, Universidad Antonio de Nebrija.
- ALBA QUIÑONES, V. (2009): "El análisis de errores en el campo del español como lengua extranjera: algunas cuestiones metodológicas", *Revista Nebrija de Lingüística Aplicada*, 5 (3), 1-16.
- ALEXOPOULOU, A. (2005): "Aproximación al tratamiento del error en la clase de E/LE desde la perspectiva del análisis de errores", *Estudios de Lingüística Aplicada*, 23 (41), 101-125.
- ALEXOPOULOU, A. (2006): "Los criterios descriptivo y etiológico en la clasificación de los errores del hablante no nativo: una nueva perspectiva", *Porta Linguarum*, 5, 17-35.
- ALEXOPOULOU, A. (2011): "El papel de la transferencia en los errores léxicos". *Revista Nebrija de Lingüística Aplicada*, 9 (5), 27-36.
- ALMEIDA FILHO, J. C. (2001): *Português para estrangeiros interface com o espanhol*. Campinas: Pontes, 2ªed.
- ANDRADE NETA, N. F. (2000): "Aprender español es fácil porque hablo portugués: Ventajas y desventajas de los brasileños para aprender español", *Cuaderno Cervantes*, Disponible en: [http://www.cuadernoscervantes.com/lc\\_portugues.html](http://www.cuadernoscervantes.com/lc_portugues.html) [consulta: 25 agosto 2006].
- APPEL, R. y MUYSKEN P. (1996): *Bilingüismo y contacto de lenguas*, Barcelona: Ariel.
- ARABSKI, J. (2006): "Language Transfer in Language Learning and Language Contact", en J. Arabski, (ed.) *Cross-linguistic Influences in the Second Language Lexicon*, Clevedon: Multilingual Matters, 12-21.

- ARCOS PAVÓN, (2009): *Análisis de errores, contrastivo e interlengua, en estudiantes brasileños de español como segunda lengua: verbos que rigen preposición y/o ausencia de ella*. Tesis doctoral, Universidad Complutense de Madrid.
- ARNAUD, P. J. L. (1984): "The lexical richness of L2 written productions and the validity of vocabulary tests", en T. Culhane, C. Klein-Braley, y D. K. Stevenson (eds.) *Practice and Problems in Language Testing*, 14-28. Colchester: University of Essex.
- ASSELIN, C. y McLAUGHLIN, A. (1992): "Les erreurs linguistiques rencontrées dans les écrits des étudiants universitaires: analyse et conséquences", *Revue de l'Association canadienne de linguistique appliquée*, 14 (1), 13-28.
- BARALO, M. (1997): "La organización del lexicón en lengua extranjera", *Revista de Filología Románica*, 14 (1), 59-73.
- BARBOSA, E. (2008): "Aspectos fonético-fonológicos en la interlengua de aprendices de español/ LE interfaz con el portugués". Congresso brasileiro de hispanistas, Minas Gerais.
- BARROS, C. S. y COSTA, E. G. M. (2010): "Elaboração de materiais didáticos para o ensino de espanhol", en C. S. Barros y E. G. M. Costa (coords.), *Espanhol ensino médio*, Coleção explorando o ensino, 16, Brasília: Ministério da Educação, Secretaria de Educação Básica.
- BASTERRECHEA MORENO, J. P. (2007): "El Instituto Cervantes en Brasil". IV CILE Congreso Internacional de la Lengua Española, Cartagena. Disponible en: [http://congresosdelalengua.es/cartagena/ponencias/seccion\\_3/37/basterrechea\\_moreno\\_juan\\_pedro\\_d.htm](http://congresosdelalengua.es/cartagena/ponencias/seccion_3/37/basterrechea_moreno_juan_pedro_d.htm) [consulta: 08 marzo 2011].
- BIBER, D. (1988): *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- BILD, E. R. y SWAIN, M. (1989): "Minority language students in a French Immersion programme: their French proficiency", *Journal of Multilingual and Multicultural Development*, 10 (3), 255-274.
- BLANCHE-BENVENISTE, C., MOTA, M. A., et al. (red.) (1997): *Eurom4. Método de ensino simultâneo das línguas românicas. Método para la enseñanza de las lenguas románicas. Metodo di insegnamento simultaneo delle lingue romanze. Méthode d'enseignement simultanée des langues romanes*. Firenze: La Nuova Italia Editrice, Scandicci & Université de Provence & Universidad de Salamanca & Università degli Studi di Roma Tre, Universidade de Lisboa.
- BOSQUE, I. (2004): *REDES Diccionario combinatorio del español contemporáneo*. Madrid: SM.
- BOU, P. (1998): "On Pragmatic Transfer", *Studies in English Language and Linguistics* 0: 5-20.

- BRAGA, J. (2010): "Calque-free lectures? Spanish cross-linguistic influence in content teaching through English", *Elia*, 10, 113-135.
- BRIONES, A. I. (2000): "Dificultades de la traducción portugués-español vistas a través de la lingüística contrastiva", en Actas del IX Congreso Brasileño de Profesores de Español, 59-68.
- BUSTOS GISBERT, J. M. (1998): "Análisis de errores: problemas de categorización", *Cuadernos de Filología Hispánica*, 16, 11-40.
- BUYSE, K. (2011): Comentarios al artículo "Los errores léxicos intralingüales semánticos en la producción escrita de los aprendientes griegos de español" de Kiriakí Palapanidi. *Revista Nebrija de Lingüística Aplicada*, 9 (5), 37- 40.
- CALVI, M. V. (2004): "Aprendizaje de lenguas afines: español e italiano", *RedELE Revista Electrónica de Didáctica del Español Lengua Extranjera*, 1.
- CALVO CAPILLA, M. C. (2010): "O efeito da língua materna na aquisição de línguas próximas", *Revista Intercâmbio dos Congressos de Humanidades*, 1-29, Disponible en: <http://unb.revistaintercambio.net.br> [consulta: 05/12/2012].
- CALVO CAPILLA, M. C. y RIDD, M. (2009): "A tradução como atividade contrastiva e de conscientização na aprendizagem de línguas próximas", *RHLA*, 8 (2), 150-169.
- CAMORLINGA, R. (1997): "A distância da proximidade - a dificuldade de aprender uma língua fácil", *Intercambio*, 6.
- CAMPOS, S. N. (2000): "Gente de Brasil: la enseñanza del español lengua extranjera para estudiantes brasileños", *Encuentro. Revista de investigación e innovación en la clase de idiomas*, 11, 183-194.
- CANCELLIER, A. (1996): *Lenguas en contacto. Italiano y español en el Río de la Plata*. Padova: Unipress.
- CANDEAS, A. (2007): "Ministerio de Educación, La enseñanza del español en Brasil: un compromiso de integración cultural". IV CILE Congreso Internacional de la Lengua Española, Cartagena, Disponible en: <http://www.congresosdelalengua.es/cartagena/ponencias> [consulta: 08 marzo 2011].
- CANTOS GOMEZ, P. (2002): "Precisamos de estatística quando temos a lingüística?", *Delta*, 18 (2), 233 - 271.
- CARCEDO GONZÁLEZ, A., (1998): "Análisis de errores léxicos del español en la interlengua de los finlandeses", *Actas del IX Congreso Internacional de la Asociación para la Enseñanza del Español como Lengua Extranjera ASELE*.
- CARROLL, J. B. (1964): "Words, meanings, y concepts", *Harvard Educational Review*, 34 (2), 178-202.

- CASTILHO, A. T. de (1989): "Para o estudo das unidades discursivas no português falado", en A.T. de Castilho (ed.), *Português Culto Falado no Brasil*. Campinas: Editora da Universidade Estadual de Campinas, 249-280.
- CASTILLEJOS-LÓPEZ, W. (2009): "Error analysis in a learner *corpus*. What are the learner's strategies?" AELINCO I Congreso Internacional de Lingüística de *Corpus* (CILC'09) Universidad de Murcia.
- CELADA, M. y RODRIGUES, F. (2005): "El español en Brasil: actualidad y memoria", *Área: Lengua y Cultura - ARI*, 31.
- CELAYA, M.L. (2006): "Lexical Transfer and Second Language Proficiency: a Longitudinal Analysis of Written Production in English as a Foreign Language", *Actas del XXIX AEDEAN*, Universidad de Jaén, España.
- CELAYA, M. L. (2007): "'I study *natus* in English': lexical transfer in CLIL and regular learners". Paper presented at the AESLA Conference, Murcia, España.
- CELAYA, M. L., PÉREZ-VIDAL, C., y TORRAS, M. R. (2001): "Matriz de criterios de medición para la determinación del perfil de competencia lingüística escrita en inglés (LE)", *Resla*, 14, 87-98.
- CELAYA, M. L. y TORRAS, M. R. (2001): "L1 influence and EFL vocabulary: do children rely more on L1 than adult learners?", *Actas del XXV AEDEAN*. Universidad de Granada, 1-14.
- CENOZ, J. (2001): "The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition", en J. Cenoz, B. Hufeisen, y U. Jessner (eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*, 8-20. Clevedon, UK: Multilingual Matters.
- CENOZ, J., HUFEISEN, B. y JESSNER U. (2003): *The multilingual lexicon*. Dordrecht: Kluwer Academic.
- CENTRO VIRTUAL CERVANTES (2012): *Diccionario de términos clave ELE*. Instituto Cervantes. Disponible en: [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/diccio\\_ele/indice.htm](http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/indice.htm) [consulta: 15/02/2012].
- CORDER, S. P. (1981): *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- CORDER, S. P. (1983): "A role for the mother tongue", en S. Gass and L. Selinker (eds.), *Language transfer in language learning*, 85-97. Rowley, MA: Newbury House.
- CORDER, S. P. (1967): "The significance of learners errors", *International Review of Applied Linguistics*, 5, 161-169.
- CORNAIRE, C. y RAYMOND, P. (1999): *La Production écrite*. Paris: Clé International.

- COVEST – COPSET (2011): *Comissão do Vestibular e Comissão de Processos Seletivos e Treinamentos*. Disponible en: <http://www.covest.com.br/default.asp> [consulta: 08/03/2011].
- CRESTI, E. (2005): "Notes on lexical strategy, structural strategies and surface clause indexes in C-ORAL-ROM spoken corpora", *C-ORAL-ROM: integrated reference corpora for spoken Romance languages*, en E. Cresti y M. Moneglia (eds.), *Studies in Corpus Linguistics*, 15, 209-256.
- CHAPELLE, C. y DUFF, P. (2003): "Some guidelines for conducting quantitative and qualitative research in TESOL", *TESOL Quarterly*, 37 (1), 157-178.
- CHAPETÓN, C. (2008): "Cross-linguistic influence in the writing of an Italian learner of English as a foreign language: An exploratory study", *Colombian Applied Linguistics Journal* 10, 50-72.
- DEWAELE, J-M. y PAVLENKO, A. (2003): "Productivity and lexical diversity in native and non-native speech: a study of cross-cultural effects", en V. J. Cook (ed.), *L2 Effects on the L1*, Clevedon: Multilingual Matters.
- DOMÍNGUEZ GONZÁLEZ, P. (2008): "Destrezas receptivas y destrezas productivas en la enseñanza del Español como lengua extranjera. Breve curso para profesores en formación", *Monográficos MarcoELE*, 6.
- DÖRNYEI, Z. (2007): *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- DUGAST, D. (1978): "Sur quoi se fonde la notion d'étendue théorique du vocabulaire?", *Le français moderne*, 1, 25-32.
- DUGAST, D. (1979): *Vocabulaire et stylistique*. 1. Théâtre et dialogue. Genève: Slatkine.
- DUGAST, D. (1980): *La statistique lexicale*. Genève: Slatkine.
- DULAY, H. y BURT M. (1974): "Natural sequences in child second language acquisition", *Language Learning*, 24, 37-53.
- DULAY, H., BURT, M. y KRASHEN S. (1982): *Language Two*. Oxford: Oxford University Press.
- DURAN, P., MALVERN, D., RICHARDS, B., y CHIPERE N. (2004): "Developmental trends in lexical diversity", *Applied Linguistics*, 25, 220-242.
- DURÃO, A. B. A. B. (2004): *Análisis de Errores en la interlengua de brasileños aprendices de español y de españoles aprendices de portugués*. 2ª edición modificada. Londrina: Eduel.
- DUŠKOVÁ, L. (1984): "Similarity - An aid or hindrance in foreign language learning?", *Folia Linguistica*, 18 (1-2), 103-115.

- ELLIS, R. (2008): *The Study of Second Language Acquisition*. (2 ed.). Oxford: Oxford University Press.
- ENGBER, C. (1995): "The Relationship of Lexical Proficiency to the Quality of ESL Compositions", *Journal of Second Language Writing*, 4 (2), 139-155.
- ERES, G. (2001): "Las variantes del español en la nueva década: ¿todavía un problema para el profesor?" *Actas del IX Seminario de Dificultades Específicas de la Enseñanza del Español a Lusohablantes*, São Paulo.
- ESPIGA, J. W. (1999). "Interferências do português no aprendizado de língua espanhola", *Comunicação II SENALE - Seminário Nacional de Linguagem e Ensino*. Pelotas, UCPel.
- FARAHANI, A. y MERAJI, S. (2011): "Cognitive Task Complexity and L2 Narrative Writing Performance", *Journal of Language Teaching and Research*, 2 (2), 445-456.
- FERNÁNDEZ GONZÁLEZ, J. (1995): *El análisis contrastivo: historia y crítica*. Valencia: Universidad de Valencia.
- FERNÁNDEZ JÓDAR, R. (2006): *Análisis de errores léxicos, morfosintácticos y gráficos en la lengua escrita de los aprendices polacos de español*. Tesis doctoral. Disponible en: <http://www.mec.es/redele/Biblioteca2007/RaulFernandez.shtml> [consulta: 27/11/2012].
- FERNÁNDEZ, S. (1997): *Interlengua y análisis de errores en el aprendizaje del español como lengua extranjera*. Madrid: Didascalía.
- FIALHO, V. R. (2005): "Proximidade entre línguas: algumas considerações sobre a aquisição o espanhol por falantes nativos de português brasileiro", *Espéculo: Revista de estudios literarios*. Universidad Complutense de Madrid.
- FIDLER, A. (2006): "Reconceptualizing fossilization in second language acquisition: a review", *Second Language Research*, 22 (3), 398-411.
- FIGUERAS, N. (2008): "El MCER, más allá de la polémica", *MarcoELE*, 7, 26-35.
- FIRTH, A., y WAGNER, J. (1997): "On discourse, communication, and (some) fundamental concepts in SLA research", *Modern Language Journal*, 81, 285-300.
- FISIAK, J. (1981): *Contrastive Linguistic and the Language Teacher*. Oxford: Oxford Pergamon Press.
- FRAUENFELDER, U., NOYAU, C., PERDUE, C. y PORQUIER, R. (1980): "Connaissance en langue étrangère", *Langages*, 57, 43-60.
- GALINDO MERINO, M. (2004): *Principales modelos de análisis de datos en la investigación sobre adquisición de segundas lenguas*. Alicante. Universitat d'Alacant.

- GALLEGO, M. T. y LLACH, M. P. (2009): "Exploring the increase of receptive vocabulary knowledge in the foreign language: A longitudinal study.", *International Journal of English Studies, IJES*, 9 (1), 113-133.
- GASS, C. y SELINKER, L. (eds.) (1983): *Language Transfer in Language Learning*. Rowley, MA: Newbury House.
- GASS, S. y SELINKER, L. (2008): *Second Language Acquisition. An introductory Course*, (3 ed.), London: Routledge.
- GILQUIN, G. (2006): "Towards an empirically grounded definition of prototypes", Poster Presentation at *Linguistic Evidence II*, Tübingen, 2-4.
- GODOY, E. (2001): "La cultura en la enseñanza del español y de las literaturas hispánicas", *Abel Anuario brasileño de estudios hispánicos*, 11, 323-341.
- GOES DE ANDRADE, O. (2006): "O que sabemos quando conhecemos uma palavra?", *VII Encontro do Círculo de Estudos Linguísticos do Sul – CELSUL*, Universidade Católica de Pelotas – UCPel, Universidade Federal de Pelotas – UFPel.
- GÓMEZ MOLINA, J. R. (1997): "La formación de palabras en clase de ELE", *REALE Revista de Estudios de Adquisición de la Lengua Española*, 8, 69-93.
- GÓMEZ PALOU, M. (2006): *Tutorial del Wordsmith 4*. Universidad de Ottawa.
- GONZÁLEZ, N. T. M. (2000): "La lengua española en Brasil. La lengua española para los brasileños", *Comunica Hipanista*, Madrid y Niterói.
- GRAUBERG, W. (1971): "An error analysis in German of first year university students", en G. Perren y J. Trim (eds.), *Applications of linguistics*. Cambridge: Cambridge University Press.
- GRECO, S. (2006): "Lenguas afines", *redELE Revista Electrónica de Didáctica de Español Lengua Extranjera*, 6.
- GUIRAUD, P. (1960): *Problèmes et méthodes de la statistique linguistique*. Dordrecht: D. Reidel.
- HAMP-LYONS, L. (1990): "Second language writing: Assessment issues", en B. Kroll (ed.), *Second language writing: Research insights for the classroom*, 69-87. Cambridge: Cambridge University Press.
- HAN, Z. (2003): "Fossilization: From Simplicity to Complexity", *International Journal of Bilingual Education and Bilingualism*, 6 (2), 95-128.
- HAN, Z. (2004): *Fossilization in Adult Second Language Acquisition*. Clevedon: Multilingual Matters Ltd.

- HAN, Z. y SELINKER, L. (2005): "Fossilization in L2 Learners", en E. Hinkel, (ed.), *Handbook of Research in Second Language Teaching and Learning*, 455–470. London: Lawrence Erlbaum Associates.
- HENRIQUES, E. R. (2000): "Intercompreensão de texto escrito por falantes nativos de português e de espanhol", *DELTA*, 16 (2), 263-295.
- HUGHES, A. y LASCARATOU, C. (1982): "Competing Criteria for Error Gravity", *Language Learning*, 36, 175-182.
- HYLTENSTAM, K. (1988): "Lexical characteristics of near-native second language learners of Swedish", *Journal of Multilingual and Multicultural Development*, 9, 67–84.
- INEP INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (2011): Disponible en: <http://portal.inep.gov.br/>, consulta [15/03/2011].
- INSTITUTO CERVANTES (2006): *Plan curricular del Instituto Cervantes. Niveles de referencia para el español*. Madrid: Instituto Cervantes.
- INSTITUTO CERVANTES (2009): *El español en el mundo. Anuario 2009*, Disponible en: [http://www.cervantes.es/lengua\\_y\\_ensenanza/publicaciones\\_espanol/espanol\\_mundo.htm](http://www.cervantes.es/lengua_y_ensenanza/publicaciones_espanol/espanol_mundo.htm) [consulta: 27/11/2012].
- ISRAEL, M. J. (2001): "Reflexões lingüísticas sobre a metodologia de ensino-aprendizagem de uma língua estrangeira: um caso do ensino do espanhol no Brasil". *Literatura lingüística*, 13, 179-190.
- JAMES, C. (1994): "Don't shoot my dodo", *IRAL - International Review of Applied Linguistics in Language Teaching*, 32 (3), pp. 179–200.
- JAMES, C. (1998): *Errors in language learning and use: Exploring error analysis*. New York: Longman.
- JARVIS, S. (1998): *Conceptual transfer in the interlanguage lexicon*. Bloomington: Indiana University Linguistics Club.
- JARVIS, S. (2000): "Methodological Rigor in the Study of Transfer: Identifying L1 Influence in the Interlanguage Lexicon", *Language Learning*, 50 (2), 245-309.
- JARVIS, S. (2002): "Short texts, best-fitting curves and new measures of lexical diversity", *Language Testing*, 19 (1), 57–84.
- JARVIS, S., GRANT, L., BIKOWSKI, D. y FERRIS, D. (2003): "Exploring multiple profiles of highly rated learner compositions", *Journal of Second Language Writing*, 12, 377–403.
- JEGERSKI, J., VANPATTEN, B. y KEATING, G. D. (2011): "Cross-linguistic variation and the acquisition of pronominal reference in L2 Spanish", *Second Language Research*, 27 (4), 481-507.



- JENSEN, J. B. (1989): "On the Mutual Intelligibility of Spanish and Portuguese", *Hispania*, 72, 4, 848-852.
- JIANG, N. (2002): "Form–meaning mapping in vocabulary acquisition in a second language", *Studies in Second Language Acquisition*, 24, 617–638.
- JORDAN, I. (1991): "Portuguese for Spanish Speakers: A Case for Contrastive Analysis", *Hispania*, 74 (3), 788-792.
- KELLERMAN, E. (1977): "Towards a characterisation of the strategy of transfer in second language learning", *Interlanguage Studies Bulletin*, 2, 58-145.
- KELLERMAN, E. (1978): "Transfer and non-transfer: where we are now", *Studies in Second Language Acquisition*, 2 (1), 37-57.
- KELLERMAN, E. (1983): "Now you see it, now you don't", en S. Gass y L. Selinker (eds.), *Language transfer in language learning*, 112-134. Rowley, MA: Newbury.
- KELLERMAN, E. (1986): "An Eye for an Eye: Crosslinguistic Constraints on the Development of the L2 Lexicon", en E. Kellerman y M. Sharwood Smith (eds.), *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon Press.
- KELLERMAN, E. (1989): "The imperfect conditional: fossilization, cross-linguistic influence and natural tendencies in a foreign language setting", en K. Hyltenstam y L. Obler (eds.), *Bilingualism across the life span*, 87–115. Cambridge: Cambridge University Press.
- KELLERMAN, E. (1995): "Crosslinguistic influence: Transfer to nowhere?", *Annual review of applied linguistics*, 15, 125-150.
- KELLERMAN, E. (2000): "What fruit can tell us about lexicosemantic transfer: a non-structural dimension to learners' perceptions of linguistic relations", en C. Munoz (ed.) *Segundas lenguas, Adquisición en el aula*. Barcelona: Ariel.
- KELLERMAN, E. (2001): "New uses for old language: cross-linguistic and cross-gestural influence in the narratives of non-speakers", en J. Cenoz, B. Hufeisen, y U. Jessner (eds.) *Cross-linguistic influence in Third Language Acquisition: Psycholinguistic Perspectives*, 170–191, Clevedon: Multilingual Matters.
- KELLERMAN, E. y SHARWOOD SMITH, M. (1986): *Crosslinguistic Influence in Second Language Acquisition*. Pergamon Press.
- KOHN, K. (1986): "The Analysis of Transfer", en E. Kellerman y M. Sharwood Smith (eds.), *Crosslinguistic Influence in Second Language Acquisition*. Oxford: Pergamon Press.
- KRASHEN, S. (1981): *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon Press.
- LADO, R. (1957): *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor, Michigan: University of Michigan.

- LARDIERE, D. (1998): "Case and tense in the "fossilized" steady state.", *Second Language Research* 14, 1–26.
- LARRINAGA, J. A. (2006): "Estudio de producciones escritas de los niveles intermedios a diferenciar: análisis formal", *4º Congresso Brasileiro de Hispanistas*, Língua espanhola, 1, 464-470.
- LARSEN-FREEMAN, D. y LONG, M. (1991): *An Introduction to Second Language Acquisition Research*. London: Longman.
- LASAGABASTER, D. (2001): "The learning of English in Finland", *Interface, Journal of Applied Linguistics*, 16, 27-44.
- LAUFER, B. (1990): "Sequence and order in the development of L2 lexis: Some evidence from lexical confusions", *Applied Linguistics*, 11, 281–296.
- LAUFER, B. (1991): "Some properties of the foreign language learner's lexicon as evidenced by lexical confusions", *International Review of Applied Linguistics*, 29, 317–330.
- LAUFER, B. (1992): "Reading in a foreign language: How does L2 lexical knowledge interact with the reader's general academic ability?", *Journal of Research in Reading*, 15, 95-103.
- LAUFER, B. (1998): "The Development of Passive and Active Vocabulary in a Second Language: Same or Different?", *Applied Linguistics*, 19 (2), 255-271.
- LAUFER, B. (2000): "Avoidance of idioms in a second language: the effect of L1-L2 degree of similarity", *Studia Linguistica*, 54 (2), 186-196.
- LAUFER, B. (2003): "The influence of L2 on L1 collocational knowledge and on L1 lexical diversity in free written expression", en V. Cook (ed.), *Effects of the second language on the first*, 19-31. Clevedon: Multilingual Matters.
- LAUFER, B. y GIRSAI N. (2008): "Form-focused Instruction in Second Language Vocabulary Learning: A Case for Contrastive Analysis and Translation", *Applied Linguistics*, 29 (4), 694–716.
- LAZARATON, A. (2005): "Quantitative Research Methods", en E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning*, 209–224. London: Lawrence Erlbaum Associates.
- LEFFA, V. J. (2000): "Aspectos externos e internos da aquisição lexical", en V. J. LEFFA (org.), *As palavras e sua companhia; o léxico na aprendizagem*, 15-44. Pelotas: ALAB e Editora da Universidade Católica de Pelotas.
- LEHMANN, M. (2007): "The lexical diversity of short texts: Exploring the productive-receptive continuum of lexical knowledge", en J. Horváth y M. Nikolov (eds.), *UPRT 2007: Empirical studies in English applied linguistics*, 293-305. Pécs: Lingua Franca Csoport.

- LEKI, I. y CARSON J. G. (1994): "Students' Perceptions of EAP Writing Instruction and Writing Needs Across the Disciplines", *TESOL QUARTERLY*, 28 (1), 81-101.
- LENARDUZZI, R. (1999): "Interferencias en el aprendizaje del español en alumnos itálofonos: el lexema verbal", en *Actas del XVIII Congreso de la a.isp.i (Associazione Ispanisti Italiani)*, Vol. II. Siena, 243-260.
- LICERAS, J. M. (1991): "¿Gramática o gramáticas del español no-nativo?", en C. Hernández et al. (ed.) *El español de América, Vol. III. Proceedings of the III International Congress of American Spanish*. Salamanca: Junta de Castilla y León, 1335-1340.
- LINNARUD, M. (1983): "On lexis: the Swedish learner and the native speaker compared", en *Cross Language Analysis and Second Language Acquisition*, 2 ed., K. Sajavaara: Jyväskylä.
- LIPSKI, J. M. (2006): "Too close for comfort? The genesis of "portuñol/portunhol".", *Selected Proceedings of the 8th Hispanic Linguistics Symposium*. T. L. Face y C. A. Klee (ed.), 1-22. Somerville, MA: Cascadilla Proceedings Project.
- LIU, S. (2001): "Studies on Transfer in Second Language Acquisition", *Guangxi Normal University Journal*, 3, 1-29.
- LOEWEN, S. y GASS, S. (2009): "The use of statistics in L2 acquisition research", *Language Teaching*, 42, 181-196.
- LONG, M. H. (2003): "Stabilization and Fossilization in Interlanguage Development", en C. Doughty y M. Long (eds.), *The Handbook of Second Language Acquisition*. Blackwell Handbooks in Linguistics, Blackwell Publishing: Malden, MA.
- LÓPEZ MORALES, H. (1993): "En torno al aprendizaje del léxico. Bases psicolingüísticas de la planificación curricular", en S. Montesa y A. Garrido (eds.), *El español como lengua extranjera: de la teoría al aula, Actas del III Congreso Internacional de la Asociación para la Enseñanza del Español como Lengua Extranjera ASELE*. Málaga: ASELE, 9-22.
- LU, X. (2011): "A Corpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development", *TESOL Quarterly*, 45 (1), 36-62.
- LLOPIS GARCÍA, R. (2009): "La instrucción gramatical en la Adquisición de Segundas Lenguas - revisión de ayer para propuestas de hoy", *redELE: Revista Electrónica de Didáctica ELE*, 16.
- MACKEY, A. y S. GASS (2005): *Second Language Research: Methodology and design*. New Jersey: Lawrence Erlbaum Associates.
- MALVERN, D., RICHARDS, B. J., CHIPERE, N. y DURAN, P. (2004): *Lexical diversity and language development: quantification and assessment*. New York: Palgrave Macmillan.

- MALVERN, D., y RICHARDS B. J. (2000): "Investigating accommodation in language proficiency interviews using a new measure of lexical diversity", *Language Testing*, 19, 85-104.
- MARANGON, G. (2011): "Estudio contrastivo del léxico en las lenguas afines. La engañosa semejanza: la traducción de los falsos amigos en español e italiano", *AnMal Electrónica*, 30, 258-292. Disponible en: [http://www.anmal.uma.es/numero30/Contraste\\_esp-ital.pdf](http://www.anmal.uma.es/numero30/Contraste_esp-ital.pdf) [consulta 27/11/2012].
- MARTÍN DEL REY, M. A. (2004): "Análisis de Errores de la interlengua de español en estudiantes italianos", Universidad de Salamanca, Máster Universitario, Publicado por Elenet.org.
- MARTÍNEZ ARBELAIZ, A. (2004): "Índices de progreso en la producción escrita de estudiantes de español en situación de inmersión", *RæI-Revistælectrónica de Lingüística Aplicada*, 3, 115-145.
- MARTÍNEZ, L. A. (2005): "Native and non-native writers' use of first person pronouns in the different sections of biology research articles in English", *Journal of Second Language Writing*, 14 (3), 174–190.
- MARTÍNEZ-CACHERO LASECA, A. (2009): "La enseñanza del español en el sistema educativo brasileño: situación y posibles actuaciones", *Lengua y Cultura - ARI* Nº 140/2009.
- MARRONE, C. S. (2005): *Português/Español: aspectos comparativos*. 2ªed., Campinas: Pontes editores.
- MARZANO, F. (2006): *Como não ficar EMBARAZADO em espanhol: Dicionário espanhol-português de fasas semelhantes*. Rio de Janeiro: Elsevier
- MASIP, V. (1999): "Fonética española para brasileiros", *Revista Gelne*, 1, 152-158.
- MCER MARCO COMÚN EUROPEO DE REFERENCIA (2002): disponible en [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/marco/](http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/) [consulta: 06/02/2012].
- MCKEE, G., MALVERN, D. y RICHARDS, B. (2000): "Measuring vocabulary diversity using dedicated software", *Literary and Linguistic Computing*, 15, 323–337.
- MCWHINNEY, B. (2000): *The CHILDES Project: Tools for analyzing talk*. New Jersey: Lawrence Erlbaum Associates.
- MEARA, P. (1984): "The study of lexis in interlanguage", en A. Davies, C. Criper y A. Howatt (eds.), *Interlanguage*. Edinburgh: Edinburgh University Press.
- MEC MINISTERIO DE EDUCACIÓN Y CIENCIA (2009): *El mundo Estudia Español* <http://www.educacion.gob.es/redele/el-mundo-estudia-esp-ol/2009.html> [consulta: 06/02/2012].
- MELLOR, A. (2011): "Essay Length, Lexical Diversity and Automatic Essay Scoring", *Memoirs of the Osaka Institute of Technology, Series B*, 55 (2), 1-14.

- MENDENHALL, W. (2007): *Introducción a la probabilidad y estadística*. 12ª ed. México: Cengage Learning Editores.
- MEO ZILIO, G. (1989): *Estudios hispanoamericanos. Temas lingüísticos*, Roma: Bulzoni.
- MERINERO, S. I. (2002): "Valoraciones subjetivas de los aprendices de español ante las semejanzas del portugués y el español", *Actas del IX Congreso Brasileño de Profesores de Español*, Brasilia: Consejería de Educación, Embajada de España, 153- 156.
- MILANI, E. M. (1999): *Gramática de Espanhol para brasileiros*. São Paulo: Editora Saraiva.
- MOHD HAYAS, K. (2006): *Errores léxicos en las composiciones de los estudiantes malayos en el aprendizaje del español como lengua extranjera*, Memoria de máster, disponible en: <http://www.mec.es/redele/Biblioteca2007/KasmaMohd.shtml> [consulta: 27/11/2012].
- MÖHLE, D. (1989): "Multilingual interaction in foreign language production", en H. W. Dechert and M. Raupach (eds.), *Interlingual Processes*, 179-194. Tübingen: Gunter Narr.
- MORENO FERNÁNDEZ, F. (2005): "El español en Brasil", en J. Sedycias (ed.), *O ensino do espanhol no Brasil – passado, presente e futuro*, 14-34. São Paulo: Parábola.
- MORILLO, M. (1998): *Datos y Cifras – La enseñanza del español en Brasil*. Brasilia: Consejería de Educación de la Embajada de España. Madrid: La Factoría.
- MOURA NEVES, M. H. (2010): "O acordo ortográfico da língua portuguesa y a meta de simplificação e unificação", *D.E.L.T.A.*, 26 (1), 87-113.
- NAKUMA, C. K. (1998): "A new theoretical account of "fossilization": implications for L2 attrition research", *International Review of Applied Linguistics*, 36 (3), 247-256.
- NARANJO HERNÁNDEZ, J. A. (2009): *Análisis de errores en la interlengua de estudiantes de español como lengua extranjera. Un estudio sobre la correlación de los errores con la comunicabilidad en la producción escrita de estudiantes de habla inglesa*, Memoria de Máster, Universitetet i Bergen.
- NATEL, T. (2002): "La proximidad entre el portugués y el español ¿facilita o dificulta del aprendizaje? El español, lengua de mestizaje y la interculturalidad". *Actas del XIII Congreso Internacional de la Asociación para la Enseñanza del Español como Lengua Extranjera*, ASELE Murcia, 825-832.
- NAVÉS, T., MIRALPEIX, I. y CELAYA M. L. (2005): "Who Transfer More ... and What? Cross-linguistic Influence in Relation to School Grade and Language Dominance in EFL", *International Journal of Multilingualism*, 2 (2), 113-134.
- NEMSER, W. (1971): "Approximative systems of foreign language learners", *International Review of Applied Linguistics*, 9, 115-123.
- ODLIN, T. (1989): *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.

- ODLIN, T. (1992): "Transferability and linguistic substrates", *Second Language Research*, 8, 171-202.
- ODLIN, T. (2003): "Cross-linguistic Influence", en C. Doughty y M. Long (eds.), *Handbook on Second Language Acquisition*, 436-486. Oxford: Blackwell.
- ODLIN, T. (2006): "Could a contrastive analysis ever be complete?", en J. Arabski (ed.) *Cross-linguistic Influences in the Second Language Lexicon*, 22-35. Clevedon: Multilingual Matters.
- OLSSON, M. (1973): "The Effects of Different Types of Errors in the Communication Situation", en J. Svartvik (ed): *Errata. Papers in Error Analysis*. GWK Gleerup, Lund.
- ORTEGA, L. (2004): "Aproximaciones cognitivo-interaccionistas al aprendizaje de segundas lenguas mediante tareas", *ELIA* 5, 15-40
- ORTEGA, L. (2009): *Understanding second language acquisition*. London: Hodder Education.
- ORTEGA, L. y IBERRI-SHEA G. (2005): "Longitudinal research in second language acquisition: Recent trends and future directions", *Annual Review of Applied Linguistics*, 25, 26-45.
- PALAPANIDI, K. (2010): "Los errores léxicos intralingües semánticos en la producción escrita de los aprendientes griegos de español", *Revista Nebrija de Lingüística Aplicada*, 9 (5), 2-21.
- PALAPANIDI, K. (2011): *Factores condicionantes de producción de errores léxicos en la lengua escrita de los aprendientes griegos de español*. Madrid: Universidad Antonio de Nebrija, tesis doctoral, s.e.
- PALAPANIDI, K. (2012): *Los errores léxicos de los aprendientes griegos de español. Datos de un corpus de producción escrita*. Editorial Académica Española.
- PENADÉS MARTÍNEZ, I. (2003): "Las clasificaciones de errores lingüísticos en el marco del análisis de errores", *Lingüística en la Red*, 1, 1-29. Disponible en [http://www.linred.es/articulos\\_pdf/LR\\_articulo\\_051120032.pdf](http://www.linred.es/articulos_pdf/LR_articulo_051120032.pdf) [Consulta en 04/12/2012].
- PÉREZ DE OBANOS, R. (2010): "La situación del español en Brasil", *Revista Electrónica de Lingüística Aplicada a la Enseñanza de Español como Lengua Extranjera*, 4, 50-74.
- POLIO, C. G. (1997): "Measures of linguistic accuracy in second language writing research", *Language Learning*, 47 (1), 101-143.
- POULISSE, N. (1990): *The Use of Compensatory Strategies by Dutch Learners of English*. Dordrecht: Foris.
- POULISSE, N. y BONGAERTS T. (1994): "First language use in second language production", *Applied Linguistics*, 15 (1), 36-57.

- PUIG, F. (2008): "El Marco Común Europeo de Referencia y la evaluación en el aula", *MarcoELE*, 7, 78-91.
- READ, J. (2000): *Assessing vocabulary*. Cambridge: CUP.
- REAL ACADEMIA ESPAÑOLA (2001): *DRAE Diccionario de la lengua española*. 22ª ed., Madrid: ESPASA Calpe.
- REID, J. (1990): "Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective", En B. Kroll (ed.), *Second language writing: Research insights for the classroom*, 191-210. New York: Cambridge University Press.
- RICHARDS, B. J., y MALVERN D. D. (1999): "The application of a new measure of lexical diversity to preschool children". *Paper presented at the 8<sup>th</sup> International Congress for the Study of Child Language*, The University of Basque Country, San Sebastian-Donostia, Spain.
- RICHARDS, J. (1971): "Error analysis and second language strategies", *Language Sciences*, 17, 12-22.
- RINGBOM, H. (1986): "Crosslinguistic Influence and the Foreign Language Learning Process", en E. Kellerman and M. Sharwood Smith (eds.), *Crosslinguistic Influence in Second Language Acquisition*. New York : Pergamon Press.
- RINGBOM, H. (1987): *The role of first language in foreign language acquisition*. Clevedon: Multilingual Matters.
- RINGBOM, H. (2001): "Lexical transfer in L3 production", en J. Cenoz, B. Hufeisen y U. Jessner (eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*, 59-68. Clevedon: Multilingual Matters.
- RINGBOM, H. (2006): "The importance of different types of similarity in transfer studies", en J. Arabski (ed.), *Cross-linguistic influences in the second language lexicon*, 36-45. Clevedon: Multilingual Matters.
- RINGBOM, H. (2007): *Cross-linguistic Similarity in Foreign Language Learning*. Clevedon: Multilingual Matters.
- RINGBOM, H. y JARVIS S. (2009): "The Importance of Cross-Linguistic Similarity in Foreign Language Learning", en M. Long y C. Doughty (eds.), *The Handbook of Language Teaching*. Wiley-Blackwell.
- RODRÍGUEZ PANIAGUA, L. R. (2001): *Análisis de errores e interlengua de aprendices árabes de español como lengua extranjera*, Memoria de máster, disponible en: <http://www.mec.es/redele/biblioteca2005/rodriguez.shtml> [consulta: 27/11/2012].
- SALSBURY, T., CROSSLEY, S., y McNAMARA, D. (2011): "Psycholinguistic word information in second language oral discourse". *Second Language Research*, 27 (3), 343-360.

- SÁNCHEZ, J. M. C. (2002): “‘Significado léxico’ y ‘significado gramatical’ en las gramáticas del español moderno”, *Boletín de la Sociedad Española de Historiografía Lingüística*, 3, 43-78.
- SÁNCHEZ IGLESIAS, J. (2003): *Errores, corrección y fosilización en la didáctica de lenguas afines: Análisis de errores en la expresión escrita de estudiantes italianos de E/ELE*, Tesis doctoral, Salamanca: Universidad de Salamanca. Disponible en: <http://gredos.usal.es/jspui/bitstream/10366/116103/1/84-7800-650-8.pdf> [consulta: 05/12/2012].
- SÁNCHEZ JIMÉNEZ, D. (2006): *Análisis de errores ortográficos de estudiantes filipinos en el aprendizaje del español como lengua extranjera*, Memoria de Máster, Ministerio de Educación y Ciencia. Biblioteca Virtual RedELE 6. Disponible en: <http://www.educacion.es/redele/Biblioteca2006/DavidSanchez.shtml> [consulta: 27/11/2012].
- SÁNCHEZ PÉREZ, A. (2001): *Gran diccionario de uso del español actual*. Madrid: SGEL.
- SANTOS GARGALLO, I. (1993): *Análisis Contrastivo, Análisis de Errores e Interlengua en el marco de la Lingüística Contrastiva*. Madrid: Síntesis.
- SANTOS GARGALLO, I. (1999): *Lingüística aplicada a la enseñanza/aprendizaje del español como lengua extranjera*. Madrid: Arco-Libros.
- SANTOS, T. (1988): “Professor's reactions to the academic writing of non-native-speaking students”, *TESOL Quarterly*, 22, 69-90.
- SANZ, C. (2000): “Bilingual education enhances third language acquisition: Evidence from Catalonia”, *Applied Psycholinguistics*, 21, 23 – 44.
- SAVILLE-TROIKE, M. (1984): “What really matters in second language learning for academic achievement?”, *TESOL Quarterly*, 18 (2), 199–219.
- SCOTT, M. (2010): *WordSmith Tools (Version 5)*. Liverpool, UK: Lexical Analysis Software. Disponible en: <http://www.lexically.net/wordsmith/version5/index.html> [consulta: 27/11/2012].
- SCHACHTER, J. (1974): “An error in error analysis”, *Language Learning*, 24, 205-214.
- SCHACHTER, J. y CELCE-MURCIA, M. (1977): “Some reservations concerning Error Analysis”. *TESOL Quarterly*, 11 (4).
- SCHMIDT, S. (1994): “Do texto ao sistema literário. Esboço de uma ciência da literatura empírica construtivista”, en H. K. Olinto (ed.), *Ciência da literatura empírica*, 53-69. Rio de Janeiro: Tempo Brasileiro.
- SCHMITZ, J. (1970): “Some Lexical Contrasts in Spanish and Portuguese: Contrastive Analysis”, *Hispania*, 53 (4), 976-979.



- SECRETARIA NACIONAL DE POLÍTICAS DE TURISMO – BRASIL (2010): Anuário Estatístico de Turismo, volume 37, ano base 2009. Disponible en: [http://www.dadosefatos.turismo.gov.br/export/sites/default/dadosefatos/anuario/downloads\\_anuario/Anuxrio\\_2010\\_-\\_Ano\\_Base\\_2009\\_\\_\\_Final\\_internet.pdf](http://www.dadosefatos.turismo.gov.br/export/sites/default/dadosefatos/anuario/downloads_anuario/Anuxrio_2010_-_Ano_Base_2009___Final_internet.pdf) [Consulta: 27/11/2012].
- SELINKER, L. (1992): *Rediscovering Interlanguage*. London: Longman.
- SELINKER, L. Y LAMENDELLA, J. (1978): "Two perspectives on fossilization in interlanguage learning", *Interlanguage Studies Bulletin*, 3, 143-191.
- SELINKER, L., (1972): "Interlanguage", *International Review of Applied Linguistics*, 10, 209-231.
- SHADISH, W. R., COOK, T. D. y CAMPBELL D. T. (2001): *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.
- SHARWOOD SMITH, M. (1979): "Strategies, language transfer and the simulation of the second language learner's mental operations", *Language Learning*, 29, 345-361.
- SHARWOOD SMITH, M. Y KELLERMAN, E. (1986): "Crosslinguistic influence in second language: an introduction", en E. Kellerman y M. Sharwood Smith (eds.), *Crosslinguistic Influence in Second Language Acquisition*, 1-9. Oxford: Pergamon.
- SILVA, E. B. (2002): "Bloqueios do aprendiz de espanhol/LE: os heterossemânticos", en CONGRESO BRASILEÑO DE HISPANISTAS, 2, San Pablo. *Scielo*. Disponible en: <http://www.scielo.org> [consulta: 8 junio 2006].
- SIMON-CEREIJIDO, G. y GUTIERREZ-CLELLEN, V. F. (2009): "A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains", *Applied Psycholinguistics*, 30, 315-337.
- SINGLETON, D. (1987): "Mother and other tongue influence on learner french: a case study", *Studies in Second Language Acquisition*, 9(3), 327-345.
- SKJAER, S. (2004): *El análisis de errores y su impacto en la comunicación en textos escritos por alumnos noruegos en su examen final de bachillerato*. Memoria de Máster, Universidad Antonio Nebrija. Disponible en: <http://www.mec.es/redele/biblioteca2005/skjaer.shtml> [consulta: 27/11/2012].
- SOTO, U. (2002): "Sobre las formas de leer: la adquisición - aprendizaje de E/LE y la relación entre sistemas ortográfico y fonológico". *Proceedings of the 2. Congreso Brasileño de Hispanistas*, 2002, São Paulo (SP, Brasil). Disponible en: <http://www.proceedings.scielo.br/> [consulta: 8 junio 2006].
- STEVENS, J. (2002): *Applied multivariate statistics for the social sciences*, (4 ed.), Mahwah, New Jersey: Erlbaum.

- STRÖMQVIST, S., JOHANSSON, V., KRIZ, S., RAGNARSDÓTTIR, H., AISENMAN, R., y RAVID, D. (2002): "Toward a crosslinguistic comparison of lexical quanta in speech and writing", *Written Language and Literacy*, 5, 45 – 67.
- TAKEUCHI, N. N. (1984): "La semejanza con la lengua materna: tropiezos para el aprendizaje del español: Ferramentas de Leitura LA", *Revista Letras*, 33.
- TORIJANO PÉREZ, J. A. (2002): *Análisis teórico práctico de los errores gramaticales en el aprendizaje del español, L-2: Expresión escrita*. Salamanca: Universidad de Salamanca.
- TORIJANO PÉREZ, J. A. (2008): El estudio de los determinantes en aprendices lusohablantes de español, *DICENDA. Cuadernos de Filología Hispánica*, 26, 235-257.
- TORRAS, M. R. (1994): "La interlengua en los primeros estadios de aprendizaje de una lengua extranjera (inglés)", *Comunicación, Lenguaje y Educación*, 24, 49-62.
- TORRAS, M. R., NAVÉS, T., CELAYA, M. L., PÉREZ VIDAL, C. (2006): "Age and IL development in writing", *Second Language Acquisition*, 156-182.
- UNIVERSIDAD DE ALCALÁ DE HENARES (2001): *SEÑAS Diccionario para la Enseñanza de la Lengua Española para Brasileños*, São Paulo: Martins Fontes.
- VALLEJO, P. M. (2008): *Estadística aplicada a las ciencias sociales*. Madrid: Universidad Pontificia Comillas.
- VANPATTEN, B. y BENATI, A. (2010): *Key terms in second language acquisition*. London: Continuum.
- VAZQUEZ, G. (1991): *Análisis de errores y aprendizaje de español / lengua extranjera*, Frankfurt: Peter Lang.
- VÁZQUEZ, G. (2010): Hacia una valoración positiva del concepto de error, monográficos marcoELE, Antología De Los Encuentros Internacionales del Español Como Lengua Extranjera. Las Navas Del Marqués, 11, 164-177.
- VENEGAS, R. (2008): "Interfaz computacional de apoyo al análisis textual: El manchador de textos", *RLA. Revista de lingüística teórica y aplicada*, 46 (2), 53-79.
- VERMEER, A. (2000): "Coming to grips with lexical richness in spontaneous speech data", *Language Testing*, 17, 65-83.
- WARDHAUGH, R. (1970): "The Contrastive Analysis Hypothesis", *TESOL Quaterly*, 4 (2), 123-130.
- WHITLEY, S. (2004): "Los errores léxicos y la adquisición de la morfología derivativa en español", *Hispania* (American Association of Teachers of Spanish and Portuguese), 87 (1), 163-172. Traducción de Marcos Cánovas.

- WOLFE-QUINTERO, K., INAGAKI, S. y KIM, H-Y. (1998): "Second language development in writing: Measures of fluency, accuracy, and complexity" (Report N. 17). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- WOODALL, B. R. (2002): "Language-switching: Using the first language while writing in a second language", *Journal of Second Language Writing*, 11, 7–28.
- WRIGHT, H. H., SILVERMAN, S. W., y NEWHOFF M. (2003): "Measures of lexical diversity in aphasia", *Aphasiology*, 17 (5), 443–452.
- YU, G. (2007): "Lexical diversity in MELAB writing and speaking task performances", *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 5, 79- 116.
- ZANÓN, J. (2007): "Psicolingüística y didáctica de las lenguas: Una aproximación histórica y conceptual", *MarcoELE, Revista de Didáctica*, 5.



# ANEXOS



# Anexo 1

Currículo del curso de Letras, UFPE.



**UNIVERSIDADE FEDERAL DE PERNAMBUCO**  
**CENTRO DE ARTES E COMUNICAÇÃO**  
**DEPARTAMENTO DE LETRAS**  
**Coordenação da Graduação do Curso de Letras**

**FLUXO CURRICULAR – LICENCIATURA EM INGLÊS OU ESPANHOL – MANHÃ**  
**INÍCIO 2010.2**

	Cód.	Nome da Disciplina	Habilitação	Nº de Créd.	C. H.	Pré-Requisito	Có-Requisito
<b>1º Período</b>	LE733	Compreensão e Produção de Texto em Língua Portuguesa	1,2	4	60		
	LE676	Leitura e Produção de Texto Acadêmico	1,2	4	60		
	PO492	Fundamentos Psicológicos da Educação	1,2	6	90		
	LE739	Inglês I: Sintaxe	1	4	60		
	LE737	Espanhol I: Fonética e Fonologia	2	4	60		
	LE735	Linguística I: Fundamentos Teóricos	1,2	4	60		
	LE736	Teoria da Literatura: Formação	1,2	4	60		
<b>2º Período</b>		Eletiva I	1,2	4	60		
		Eletiva II	1	4	60		
	LE760	Latim I: Morfologia I	2	4	60		
	LE804	Inglês II: Sintaxe II	1	4	60		
	LE749	Espanhol II: Morfologia	2	4	60	LE737	
	LE742	Linguística II: Teorias Linguísticas	1,2	4	60	LE735	
	SF451	Fundamentos da Educação	1,2	4	60		
LE744	Teoria da Literatura II: Poesia	1,2	4	60	LE736		
<b>3º Período</b>		Eletiva III	1	4	60		
	LE741	Cultura Brasileira I	1,2	4	60		
	TE707	Didática	1,2	4	60		
	LE802	Fonética e Fonologia da Língua Inglesa	1	4	60		
	LE761	Latim II: Morfologia II	2	4	60	LE760	
	LE832	Inglês III: Sintaxe III	1	4	60	LE804	
	LE750	Espanhol III: História da Língua	2	4	60	LE749	
	LE743	Linguística III: Linguística Aplicada	1,2	4	60	LE742	
LE766	Literatura Latina: Antiguidade Clássica	2	4	60	LE760		
<b>4º Período</b>		Eletiva IV	2	4	60		
	LE806	Inglês IV: Morfosintaxe I	1	4	60	LE832	
	LE751	Espanhol IV: Morfosintaxe I	2	4	60	LE750	
	LE811	Linguística Aplicada a Língua Inglesa I: Aquisição da Língua	1	4	60		
	LE817	Metodologia de Ensino de Língua Inglesa I	1	4	60	TE707, LE832, PO492, SF451	
	LE813	Literatura de Língua Inglesa I: Primórdios ao Renascimento	2	4	60	LE832	
	LE762	Literatura de Língua Espanhola I: Formação	2	4	60	LE750	
	LE797	Cultura dos Povos de Língua Inglesa	2	4	60		
	LE748	Cultura dos Povos de Língua Espanhola	2	4	60		
	LE767	Metodologia de Ensino de Língua Espanhola I	2	4	60	TE707, LE750, PO492, SF451	
AP493	Políticas Educacionais, Organização e Funcionamento da Escola Básica	1,2	4	60			

Habilitações: **1** – Licenciatura em Inglês

**2** – Licenciatura em Espanhol

5º Período		Eletiva III	2	4	60	
	LE818	Metodologia de Ensino de Língua Inglesa II	1	4	60	LE817
	LE768	Metodologia de Ensino de Língua Espanhola II	2	4	60	LE767
	LE763	Literatura de Língua Espanhola II: Período Colonial	2	4	60	LE762
	LE814	Literatura de Língua Inglesa II: Séculos XVII, XVIII e XIX	1	4	60	LE813
	LE716	Introdução a Línguas	1,2	4	60	
	LE807	Inglês V: Morfossintaxe II	1	4	60	LE806
	LE812	Linguística Aplicada a Língua Inglesa II	2	4	60	
	LE798	Estágio Curricular em Inglês I	1	4	90	LE817
	LE756	Estágio Curricular em Espanhol I	2	4	90	LE767
LE752	Espanhol V: Morfossintaxe II	2	4	60	LE751	
6º Período		Eletiva IV	1,2	4	60	
	PO493	Avaliação da Aprendizagem	1,2	4	60	
	LE815	Literatura de Língua Inglesa III: Séculos XIX e XX - Modernismo	1	4	60	LE814
	LE764	Literatura de Língua Espanhola III: Séculos XIX e XX	2	4	60	LE763
	LE799	Estágio Curricular em Inglês II	1	4	90	LE818
	LE757	Estágio Curricular em Espanhol II	2	4	90	LE768
	LE808	Inglês VI: Semântica - Pragmática	1	4	60	LE807
	LE753	Espanhol VI: Construções Sintáticas	2	4	60	LE752
	LE819	Metodologia de Ensino de Língua Inglesa III	1	4	60	LE818
	LE769	Metodologia de Ensino de Língua Espanhola III	2	4	60	LE768
7º Período	LE745	Trabalho de Conclusão de Curso I	1,2	2	30	
	LE809	Inglês VII: Estudos Avançados	1	4	60	LE808
	LE800	Estágio Curricular em Inglês III	1	5	135	LE819
	LE758	Estágio Curricular em Espanhol III	2	5	135	LE769
	LE754	Espanhol VII: Semântica	2	4	60	LE753
	AP492	Gestão Educacional e Gestão Escolar	1,2	4	60	
	LE816	Literatura de Língua Inglesa IV: Pós-Modernismo	1	4	60	LE815
	LE765	Literatura de Língua Espanhola IV: Contemporânea	2	4	60	LE764
	LE820	Metodologia de Ensino de Língua Inglesa IV	1	4	60	LE819
	LE770	Metodologia de Ensino de Língua Espanhola IV	2	3	60	LE769
8º Período	LE796	Compreensão e Produção de Texto em Língua Inglesa	1	4	60	LE809
	LE747	Compreensão e Produção de Texto em Língua Espanhola	2	4	60	LE754
	LE746	Trabalho de Conclusão de Curso II	1,2	1	30	LE745
	LE801	Estágio Curricular em Inglês IV	1	4	90	LE820
	LE759	Estágio Curricular em Espanhol IV	2	4	90	LE770
	LE810	Inglês VIII: Escrita da Língua Inglesa	1	4	60	LE809
	LE821	Metodologia de Ensino de Língua Inglesa V	1	3	45	LE820
	LE771	Metodologia de Ensino de Língua Espanhola V	2	3	45	LE770
	LE755	Espanhol VIII: Lexicografia	2	4	60	LE754



	Cód.	Nome da Disciplina	Habilitação	Nº de Créd.	C. H.	Pré-Requisito	Có-Requisito
1º Período	AR005	História das Artes	CG	4	60	---	---
	BU030	Metodologia do Estado	CG	2	30	---	---
	ED001	Educação Física	Opcativa (Obs. 1)	2	30	---	---
	LE251	Teoria da Literatura 1	CG	4	60	---	---
	LE005	Língua Portuguesa 3	CG	6	90	---	---
	LE013	Língua Francesa 3 ou	CG	6	90	---	---
	LE023	Língua Inglesa 3 ou	CG	6	90	---	---
2º Período	LE380	Língua Espanhola 3 A	CG	6	90	---	---
	LE004	Língua Portuguesa 4	TC	4	60	LE005	---
	LE248	Língua Latina 1 A	TC	4	60	---	---
	LE015	Língua Francesa 4 B	B, 3, 4	4	60	LE013	---
	LE025	Língua Inglesa 4 B	C, 5, 6	4	60	LE023	---
	LE381	Língua Espanhola 4 A	D	4	60	LE380	---
	LE552	Teoria da Literatura 2 A	A, 1, 2	3	45	LE251	---
	LE554	Teoria da Literatura 4	B, C, D, 3, 4, 5, 6	3	45	LE251	---
	LE500	Cultura Brasileira	A, 1, 2, 3, 4, 5, 6	3	45	---	---
3º Período	LE001	Literatura Portuguesa 1	TC	3	45	LE251	---
	LE351	Linguística 1	TC	4	60	---	---
	LE265	Língua Portuguesa 5 A	TC	4	60	LE004	---
	LE249	Língua Latina 2 A	TC	3	45	LE248	---
	LE211	Língua Francesa 5	B, 3, 4	4	60	LE015	---
	LE224	Língua Inglesa 5	C, 5, 6	4	60	LE025	---
	LE382	Língua Espanhola 5	D	4	60	LE381	---
	LE553	Teoria da Literatura 3	A, 1, 2	3	45	LE251	---
	LE302	Literatura Portuguesa 2	TC	3	45	LE251	---
	LE601	Literatura Brasileira 1 A	TC	4	60	LE251	---
4º Período	LE352	Linguística 2	TC	3	45	LE351	---
	LE605	Cultura Brasileira 2	A, 1, 2	3	45	---	---
	LE266	Língua Portuguesa 6 A	TC	4	60	LE249	---
	LE448	Língua Latina 3	TC	3	45	LE249	---
	LE212	Língua Francesa 6	B, 3, 4	4	60	LE211	---
	LE225	Língua Inglesa 6	C, 5, 6	4	60	LE224	---
	LE383	Língua Espanhola 6	D	4	60	LE382	---
	LE353	Linguística 3	TC	3	45	LE351	---
	LE309	Literatura Portuguesa 3 A	TC	3	45	LE251	---
	LE602	Literatura Brasileira 2 A	TC	4	60	LE251	---
LE611	Literatura de Língua Francesa 1	B, 3, 4	4	60	LE251, LE211	---	
LE621	Literatura de Língua Inglesa 1	C, 5, 6	4	60	LE251, LE224	---	
LE386	Literatura de Língua Espanhola 1	D	4	60	LE251, LE382	---	

	Habilitações	Código	Carga Horária Plena	Carga Horária de Eletivas
A	Licenciatura em Língua Portuguesa	1226	2.400 horas = 164 créd.	180 horas = 12 créd.
B	Licenciatura em Língua Portuguesa e Língua Francesa	1227	2.580 horas = 172 créd.	
C	Licenciatura em Língua Portuguesa e Língua Inglesa	1228	2.580 horas = 172 créd.	
D	Licenciatura em Língua Portuguesa e Língua Espanhola	1229	2.580 horas = 172 créd.	
1	Bacharelado em Língua Portuguesa (Pesquisador)	1219	2.340 horas = 156 créd.	180 horas = 12 créd.
2	Bacharelado em Língua Portuguesa (Crítico Literário)	1220	2.370 horas = 158 créd.	180 horas = 12 créd.
3	Bacharelado em Língua Francesa (Pesquisador)	1221	2.445 horas = 163 créd.	90 horas = 06 créd.
4	Bacharelado em Língua Francesa (Tradutor)	1222	2.445 horas = 163 créd.	90 horas = 06 créd.
5	Bacharelado em Língua Inglesa (Pesquisador)	1223	2.460 horas = 164 créd.	90 horas = 06 créd.
6	Bacharelado em Língua Inglesa (Tradutor)	1224	2.460 horas = 164 créd.	90 horas = 06 créd.

Legenda: TC = Tronco Comum e CG = Ciclo Geral

#### ORIENTAÇÃO ACADÊMICA:

Obs.1: O aluno deverá cursar uma disciplina eletiva para compensar a carga horária de Educação Física;  
 Obs.2: Os alunos de Licenciatura, com ingresso a partir de 97.1, cursarão 150 horas de Prática de Ensino;  
 Obs.3: Não existe abono de falta. O aluno deverá requerer acompanhamento especial;

Obs.4: Toda solicitação deve ser processada através de Requerimento Geral, na Coordenação do Curso;  
 Obs.5: Verificação de Aprendizagem: Resolução nº 04/94 – CCEPE;  
 Obs.6: 1(un) crédito equivale a 15 horas/aula.

#### COORDENAÇÃO DE LETRAS

FONE: 2126-8307

Horário de Atendimento:  
2ª a 6ª das 08:00 às 17:00

E-mail: coordletrasufpe@yahoo.com.br

#### ESTRUTURA ORGANIZACIONAL

- COORDENADORAS: Profª Siane Góis / Profª Stella Telles
- VICE – COORD.: Prof. Francisco Costa / Profª. Ana Lima
- SECRETÁRIA: Mércia Souza
- ASSISTENTE: Luciana Simões

5º Período	LE213	Língua Francesa 7	B, 3, 4	4	60	LE212	-----
	LE226	Língua Inglesa 7	C, 5, 6	4	60	LE225	-----
	LE384	Língua Espanhola 7	D	4	60	LE383	-----
	LE604	Literatura Portuguesa 4 A	A, 1, 2	4	60	LE251	-----
	LE605	Literatura Brasileira 3 A	TC	4	60	LE251	-----
	LE612	Literaturas de Língua Francesa 2	B, 3, 4	4	60	LE251, LE211	-----
	LE622	Literaturas de Língua Inglesa 2	C, 5, 6	4	60	LE251, LE224	-----
	LE387	Literaturas de Língua Espanhola 2	D	4	60	LE251, LE382	-----
	LE358	Metodologia da Pesquisa 1	2, 3, 5	3	45	-----	-----
	SF451	Fundamentos da Educação	A, B, C, D	4	60	-----	-----
TE207	Didática 1	A, B, C, D	4	60	-----	SF451	
6º Período	LE268	Língua Portuguesa 8 A	TC	4	60	LE265	-----
	LE442	Literatura Latina 2 A	A, 1, 2	4	60	LE441	-----
	LE617	Língua Francesa 8 A	B, 3, 4	4	60	LE215	-----
	LE627	Língua Inglesa 8 A	C, 5, 6	4	60	LE226	-----
	LE385	Língua Espanhola 8	D	4	60	LE384	-----
	LE504	Literatura Brasileira 4	A, 1, 2	3	45	LE251	-----
	LE613	Literaturas de Língua Francesa 3	B, 3, 4	4	60	LE251, LE211	-----
	LE623	Literaturas de Língua Inglesa 3	C, 5, 6	4	60	LE251, LE224	-----
	LE388	Literaturas de Língua Espanhola 3	D	4	60	LE251, LE382	-----
	LE359	Metodologia da Pesquisa 2	1	4	60	-----	-----
PO403	Psicologia da Educação 6	A, B, C, D	4	60	-----	-----	
PO404	Psicologia da Educação 7	A, B, C, D	4	60	-----	-----	
AP493	Políticas Educacionais - Organização e Funcionamento da Escola Básica	A, B, C, D	4	60	-----	SF451	
7º Período	LE609	Língua Portuguesa 9 A	3, 4, 5, 6	4	60	LE268	-----
	LE402	Língua Portuguesa 10	2	5	75	LE268	-----
	LE410	Língua Francesa 8	3, 4	5	75	LE617	-----
	LE420	Língua Inglesa 8	5, 6	5	75	LE627	-----
	LE505	Literatura Brasileira 5	7	3	45	LE553	-----
	LE614	Literaturas de Língua Francesa 4	3, 4	4	60	LE613	-----
	LE624	Literaturas de Língua Inglesa 4	5, 6	4	60	LE623	-----
	LE555	Teoria da Literatura 5	3	4	60	LE553	-----
	LE404	Teoria e Prática de Pesquisa em LP	1	4	60	LE268	-----
	LE610	Teoria e Técnica de Tradução de L. F. 1	4	4	60	LE213	-----
	LE620	Teoria e Técnica de Tradução de L. I. 1	6	4	60	LE226	-----
	LE556	Crítica Literária 1	2	4	60	LE552	-----
	LE454	Linguística 4 A	1	4	60	LE353	-----
	LE419	Cultura dos Povos de Língua Francesa	3, 4	3	45	-----	-----
	LE620	Cultura dos Povos de Língua Inglesa	5, 6	3	45	-----	-----
	TE632/480	Prática de Ensino de Português 1	A, B, C, D	10	150	-----	SF451, TE207
	TE634/482	Prática de Ensino de Língua Francesa 1	B	10	150	-----	SF451, TE207
TE636/484	Prática de Ensino de Inglês 1	C	10	150	-----	SF451, TE207	
TE638/488	Prática de Ensino de Espanhol 1	D	10	150	-----	SF451, TE207	
8º Período	LE403	Língua Portuguesa 11	1	5	75	LE268	-----
	LE411	Língua Francesa 9	3, 4	5	75	LE410	-----
	LE421	Língua Inglesa 9	5, 6	5	75	LE420	-----
	LE616	Teoria e Técnica de Tradução de L. F. 2	4	4	60	LE610	-----
	LE626	Teoria e Técnica de Tradução de L. I. 2	6	4	60	LE620	-----
	LE615	Literaturas de Língua Francesa 5	3	4	60	LE614	-----
	LE625	Literaturas de Língua Inglesa 5	5	4	60	LE624	-----
	LE455	Linguística 5 A	1	4	60	LE454	-----
	LE450	Tópicos de Linguística	3, 4, 5, 6	3	45	LE353 e LE610 ou LE353 e LE620	-----
	LE556	Teoria da Literatura 6	2	4	60	LE555	-----
	LE557	Crítica Literária 2	2	4	60	LE552	-----
	LE355	Análise de Textos Literários	2	4	60	LE353	-----
	LE413	Tradução Orientada	4, 6	3	45	LE610 ou LE620	-----
	LE553	Monografia 3	1	3	45	-----	-----
	LE554	Monografia 4	2	3	45	-----	-----
	LE555	Monografia 5	3, 5	3	45	-----	-----
	TE633/481	Prática de Ensino de Português 2	A, B, C, D	10	150	SF451, TE207, TE632	PO403, PO404, AP493
TE635/483	Prática de Ensino de Francês 2	B	10	150	SF451, TE207, TE634	PO403, PO404, AP493	
TE637/485	Prática de Ensino de Inglês 2	C	10	150	SF451, TE207, TE636	PO403, PO404, AP493	
TE639/488	Prática de Ensino de Espanhol 2	D	10	150	SF451, TE207, TE638	PO403, PO404, AP493	

## Anexo 2

Cuestionario de datos personales.

Universidade Federal de Pernambuco - Centro de Artes e Comunicação - Departamento de Letras

Asignatura: \_\_\_\_\_ Profesor/a: \_\_\_\_\_ Fecha: \_\_\_\_\_

Por favor, rellena esta ficha con tus datos personales:

1. Nombre: \_\_\_\_\_

2. Apellidos: \_\_\_\_\_

3. Lugar de nacimiento: \_\_\_\_\_

4. Nacionalidad: \_\_\_\_\_

5. Sexo: hombre      mujer      6. Edad: \_\_\_\_\_

7. Correo electrónico: \_\_\_\_\_

8. Lenguas que hablas a diario/regularmente a parte del portugués: \_\_\_\_\_

9. Profesión de los padres: \_\_\_\_\_

10. ¿En qué año hiciste el “vestibular”? \_\_\_\_\_

11. a. ¿Cuál es el idioma extranjero que elegiste para el “vestibular”?

inglés      español      francés      alemán

b. ¿Por qué? \_\_\_\_\_

12. ¿Has estudiado español antes de empezar el curso de Letras?      Sí      No

En caso afirmativo, especifica:      ¿Dónde? \_\_\_\_\_

¿Durante cuánto tiempo? \_\_\_\_\_

Horas por semana: \_\_\_\_\_

13. ¿Con qué edad has comenzado a estudiar español? \_\_\_\_\_

14. ¿Estudiaste / estás estudiando alguna lengua extranjera a parte del español?

Sí      No

En caso afirmativo, especifica:      ¿Qué lengua? \_\_\_\_\_

¿Dónde? \_\_\_\_\_

¿Durante cuánto tiempo? \_\_\_\_\_

Horas por semana: \_\_\_\_\_

15. ¿Has estado alguna vez en algún país de lengua extranjera?      Sí      No

En caso afirmativo, especifica:

país (1): \_\_\_\_\_

país (2): \_\_\_\_\_

país (3): \_\_\_\_\_

edad: \_\_\_\_\_

edad: \_\_\_\_\_

edad: \_\_\_\_\_

duración: \_\_\_\_\_

duración: \_\_\_\_\_

duración: \_\_\_\_\_

motivo del viaje: \_\_\_\_\_

motivo del viaje: \_\_\_\_\_

motivo del viaje: \_\_\_\_\_

















## Anexo 4

Cuestionario sobre la percepción de la proximidad lingüística.

Universidade Federal de Pernambuco - Centro de Artes e Comunicação - Departamento de Letras

Asignatura: \_\_\_\_\_ Profesor/a: \_\_\_\_\_ Fecha: \_\_\_\_\_

1. Nombre: \_\_\_\_\_
2. Apellidos: \_\_\_\_\_
3. a. En general, ¿te parece fácil escribir en español?      Sí                      No  
b. ¿Por qué? \_\_\_\_\_  
\_\_\_\_\_
4. ¿Cuánto tiempo, aproximadamente, has utilizado para escribir el texto? \_\_\_\_\_
5. a. ¿Has tenido alguna dificultad?                      Sí                      No  
b. En caso afirmativo, ¿de qué tipo? \_\_\_\_\_  
\_\_\_\_\_
6. a. Si tuvieras que escribir el mismo texto en portugués, ¿sería diferente? Sí                      No  
b. En caso afirmativo, ¿en qué aspectos? \_\_\_\_\_  
\_\_\_\_\_
7. ¿Cómo valoras tus habilidades en lengua española?

**Insuficiente:** Tienes muchas dificultades; desconoces palabras, estructuras de la lengua.

**Regular:** Sientes dificultad, pero eres capaz de comprender / expresar algunas palabras / frases.

**Bueno:** Tienes un buen conocimiento de la lengua, aunque cometes errores que no impiden la comunicación.

**Muy bueno:** No tienes dificultad para comprender una película, leer un libro, participar de una conversación, redactar una carta para solicitar información, aunque todavía tienes dudas en algunos casos y cometes errores que no impiden la comunicación.

**Excelente:** Lo haces como en tu lengua materna, no tienes ningún tipo de dificultad.

	Insuficiente	Regular	Bueno	Muy bueno	Excelente
a. Comprensión oral					
b. Expresión oral					
c. Comprensión escrita					
d. Expresión escrita					

8. a. ¿Utilizas el español fuera del aula?                      Sí                      No  
b. ¿En qué situaciones / con quién? \_\_\_\_\_  
\_\_\_\_\_

9. ¡Muchas gracias!

Tu colaboración es muy importante.

Si quieres saber más, si tienes alguna duda, comentario o sugerencia, consulta la página

[www.geocities.com/c\\_b\\_rocha/1.html](http://www.geocities.com/c_b_rocha/1.html)



## Anexo 5

Ejemplo de redacción codificada (E3\_01\_1.1).

Asignatura: Lengua Española 3/Redacción 1: El español y tú

<FW><POS>mi <CW><SUS>relación <FW><PREP>con <FW><ART>la  
<CW><SUS>lengua <CW><VER><PRES>es  
<CW><ADV><F><INTRA><ERR>mucho <CW><ADJ>reciente .

<CW><VER><F><INTRA><ERR><IND>empezé <FW><PREP>a  
<CW><VER><F><INTRA><ERR><INF>estudié <FW><ART>el <CW><SUS>idioma  
<CW><VER><M><AMB><ERR><PRES>a <CW><ADV>más <FW><CONJ>o  
<CW><ADV>menos <FW><NUM>tres <CW><SUS>años <FW><PREP>en  
<FW><ART>la <CW><SUS><C><INTER><ERR>escuela , <FW><PREP>en  
<FW><ART><ERR>lo <FW><NUM>segundo <CW><SUS>año  
<FW><PREP><ERR>del <CW><SUS><F><INTER><ERR><PORT>ensino  
<CW><ADJ><F><INTRA><ERR>medio .

<CW><ADV>tal <CW><ADV>vez <FW><PREP>por <FW><DEM>eso  
<CW><VER><PRES>tengo <FW><ART>un <CW><SUS>poco <FW><PREP>de  
<CW><SUS><F><INTER><ERR>dificultad <FW><PREP>con  
<CW><SUS><F><INTER><ERR>vocábolos , <CW><SUS>palabras  
<FW><CONJ>y <CW><SUS><F><INTRA><ERR>construcción <FW><PREP>de  
<CW><SUS>frases , <CW><VER><F><INTER><ERR><GER><PORT>tendo  
<FW><PREP>en <CW><SUS>vista <FW><CONJ>que <CW><VER><IND>fue  
<FW><ART>un <CW><SUS>contacto <CW><ADV>muy <CW><ADJ>estrecho .

<FW><PRO>yo <CW><VER><F><INTER><ERR><PRES><PORT>penso  
<FW><PREP>en <CW><VER><F><INTRA><ERR><INF>empezé <FW><ART>un  
<CW><SUS>curso <FW><PREP>en <FW><ART><ERR>la  
<CW><SUS><F><INTER><ERR>ferias <FW><PREP>para  
<CW><VER><INF>mejorar <FW><CONJ>y <CW><ADV>también  
<FW><PREP>con <FW><ART>la <CW><SUS>finalidad <FW><PREP>de  
<CW><VER><INF>ayudarme <CW><ADV>ya <FW><CONJ>que  
<CW><VER><PRES>estoy <CW><VER><C><INTER><ERR><GER>precisando  
<CW><VER><INF>tener <CW><SUS>base <FW><PREP>para <FW><ART>las  
<CW><SUS>clases <FW><PREP>de <FW><ART>la <CW><SUS>universidad .

<FW><POS>mi <CW><SUS>relación <CW><ADV>no <CW><VER><PRES>es  
<CW><ADJ>mala , <FW><CONJ>pero <CW><ADV>no  
<CW><VER><F><INTER><ERR><PRES>és <FW><PREP>de  
<CW><ADJ><F><INTER><ERR><PORT>imenso <CW><SUS>amor .

<CW><VER><PRES>creo <FW><CONJ><ERR>qué <CW><ADV>solamente  
<CW><VER><GER>estudiando <FW><PREP>con  
<CW><SUS><F><INTER><ERR><PORT>afinco , <CW><VER><FUT>podré  
<CW><VER><INF>desarrollar <FW><PREP>con  
<CW><ADV><F><INTRA><ERR>mas <CW><SUS>seguridad <FW><DEM>esta  
<CW><SUS>lengua , <FW><CONJ>que  
<CW><VER><F><INTER><ERR><PRES><PORT>confeso <FW><PRO>me  
<CW><VER><PRES>encanta .



## Anexo 6

Pruebas de normalidad (Shapiro-Wilkinson).

Tabla 31. Test de normalidad: variables palabras total y errores total por asignatura.

Variables	Asignatura	Shapiro-Wilk			Distribución normal
		Estadística	gl	Sig.	
Palabras total ( <i>tokens</i> )	Lengua Española 3	,928	48	,006	*
	Lengua Española 4	,971	75	,078	
	Lengua Española 6	,961	63	,046	
	Lengua Española 7	,965	21	,631	
	Lengua Española 8	,858	24	,003	
	Prácticas de Enseñanza I	,954	27	,262	
	Prácticas de Enseñanza II	,990	107	,644	
Errores total ( <i>tokens</i> )	Lengua Española 3	,757	48	,000	*
	Lengua Española 4	,938	75	,001	
	Lengua Española 6	,978	63	,329	
	Lengua Española 7	,959	21	,501	
	Lengua Española 8	,958	24	,390	
	Prácticas de Enseñanza I	,963	27	,438	
	Prácticas de Enseñanza II	,906	107	,000	

\* valor-p > 0,05=distribución normal; gl=grados de libertad; sig.=significación.

Tabla 32. Test de normalidad: variable palabras total.

Variables	Bloques	Shapiro-Wilk			Distribución normal
		Estadística	gl	Sig.	
<i>types</i> total	1	,969	123	,006	*
	2	,990	108	,623	
	3	,993	134	,795	
<i>tokens</i> total	1	,963	123	,002	*
	2	,975	108	,036	
	3	,989	134	,335	
<i>types</i> léxicos	1	,964	123	,002	*
	2	,982	108	,158	
	3	,992	134	,681	
<i>tokens</i> léxicos	1	,961	123	,001	*
	2	,979	108	,090	
	3	,991	134	,543	
<i>types</i> funcionales	1	,966	123	,004	*
	2	,985	108	,292	
	3	,990	134	,443	
<i>tokens</i> funcionales	1	,965	123	,003	*
	2	,965	108	,006	
	3	,985	134	,158	

\* valor-p > 0,05=distribución normal; gl=grados de libertad; sig.=significación.

Tabla 33. Test de normalidad: variable errores.

Variables	Bloques	Shapiro-Wilk			Distribución normal
		Estadística	gl	Sig.	
<i>types</i> errores total	1	,909	123	,000	
	2	,965	108	,007	
	3	,929	134	,000	
<i>tokens</i> errores total	1	,874	123	,000	
	2	,963	108	,004	
	3	,918	134	,000	
<i>types</i> errores léxicos	1	,932	123	,000	
	2	,961	108	,003	
	3	,895	134	,000	
<i>tokens</i> errores léxicos	1	,904	123	,000	
	2	,960	108	,003	
	3	,883	134	,000	
<i>types</i> errores funcionales	1	,883	123	,000	
	2	,938	108	,000	
	3	,895	134	,000	
<i>tokens</i> errores funcionales	1	,826	123	,000	
	2	,932	108	,000	
	3	,884	134	,000	

valor-p >0,05=distribución normal; gl=grados de libertad; sig.=significación.



Tabla 34. Test de normalidad: variable categoría de palabra léxica.

Variables	Bloques	Shapiro-Wilk			Distribución normal
		Estadística	gl	Sig.	
<i>types</i> sustantivo	1	,902	123	,000	
	2	,946	108	,000	
	3	,966	134	,002	
<i>tokens</i> sustantivo	1	,916	123	,000	
	2	,947	108	,000	
	3	,976	134	,018	
<i>types</i> adjetivo	1	,930	123	,000	
	2	,957	108	,001	
	3	,956	134	,000	
<i>tokens</i> adjetivo	1	,915	123	,000	
	2	,974	108	,033	
	3	,957	134	,000	
<i>types</i> verbo	1	,958	123	,001	
	2	,968	108	,011	
	3	,975	134	,015	
<i>tokens</i> verbo	1	,950	123	,000	
	2	,971	108	,019	
	3	,978	134	,031	
<i>types</i> adverbio	1	,965	123	,003	
	2	,975	108	,038	
	3	,971	134	,006	
<i>tokens</i> adverbio	1	,929	123	,000	
	2	,974	108	,033	
	3	,972	134	,007	

valor-p >0,05=distribución normal; gl=grados de libertad; sig.=significación.

Tabla 35. Test de normalidad: variable categoría de palabra léxica con errores.

Variable	Bloque	Shapiro-Wilk			Distribución normal
		Estadística	gl	Sig.	
<i>types</i> errores sustantivo	1	,914	123	,000	
	2	,899	108	,000	
	3	,821	134	,000	
<i>tokens</i> errores sustantivo	1	,919	123	,000	
	2	,888	108	,000	
	3	,778	134	,000	
<i>types</i> errores adjetivo	1	,810	122	,000	
	2	,858	108	,000	
	3	,784	134	,000	
<i>tokens</i> errores adjetivo	1	,809	123	,000	
	2	,848	108	,000	
	3	,784	134	,000	
<i>types</i> errores verbo	1	,938	123	,000	
	2	,919	108	,000	
	3	,893	134	,000	
<i>tokens</i> errores verbo	1	,911	123	,000	
	2	,916	108	,000	
	3	,904	134	,000	
<i>types</i> errores adverbio	1	,802	123	,000	
	2	,768	108	,000	
	3	,721	134	,000	
<i>tokens</i> errores adverbio	1	,770	123	,000	
	2	,764	108	,000	
	3	,704	134	,000	

valor-p >0,05=distribución normal; gl=grados de libertad; sig.=significación.

Tabla 36. Test de normalidad: variable tipo de error léxico.

Variables	Bloques	Shapiro-Wilk			Distribución normal
		Estadística	gl	Sig.	
errores formales	1	,902	123	,000	
	2	,955	108	,001	
	3	,865	134	,000	
errores semánticos	1	,852	123	,000	
	2	,716	108	,000	
	3	,814	134	,000	
errores mixtos	1	,832	123	,000	
	2	,788	108	,000	
	3	,698	134	,000	

valor-p >0,05=distribución normal; gl=grados de libertad; sig.=significación.

Tabla 37. Test de normalidad: variable origen del error léxico.

Variables	Bloques	Shapiro-Wilk			Distribución normal
		Estadística	gl	Sig.	
Errores interlingüísticos	1	,862	123	,000	
	2	,912	108	,000	
	3	,830	134	,000	
Errores intralingüísticos	1	,962	123	,002	
	2	,939	108	,000	
	3	,920	134	,000	
Errores ambiguos	1	,699	123	,000	
	2	,668	108	,000	
	3	,514	134	,000	

valor-p >0,05=distribución normal; gl=grados de libertad; sig.=significación.

Tabla 38. Test de normalidad: variable la riqueza léxica.

	Bloques	Shapiro-Wilk			Distribución normal
		Statistic	df	Sig.	
Uber U	1	,910	123	,000	
	2	,895	108	,000	
	3	,885	134	,000	
Uber U sin errores	1	,963	123	,002	
	2	,918	108	,000	
	3	,871	134	,000	

valor-p >0,05=distribución normal; gl=grados de libertad; sig.=significación.



## Anexo 7

Pruebas no paramétricas (Kruskal-Wallis).

Tabla 39. Test de Kruskal-Wallis: fluidez.

Palabras total ( <i>tokens</i> )												
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p	
Bq1	123	101,00	99	1218,50	3,14	34,90	33,00	228,00	137,22	36,973	,000	
Bq2	108	122,93	117	1597,22	3,84	39,96	39,00	245,00	195,49			
Bq3	134	132,64	132	2242,24	4,09	47,35	24,00	253,00	214,96			

Tabla 40. Test de Kruskal-Wallis: diversidad léxica.

Panel A: Uber U												
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p	
Bq1	123	26,9144	26	40,007	,57032	6,32513	15,77	55,71	195,10	5,798	,05	
Bq2	108	25,1639	24	27,778	,50715	5,27046	15,82	43,98	163,00			
Bq3	134	26,4372	25	36,913	,52485	6,07557	16,15	56,98	188,01			
Panel B: Uber U sin errores												
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p	
Bq1	123	21,1922	21	22,156	,42442	4,70705	12,73	37,30	162,82	19,229	,000	
Bq2	108	21,4553	20	22,848	,45995	4,77999	12,11	37,37	166,61			
Bq3	134	23,6401	23	31,368	,48383	5,60076	13,52	55,30	214,74			

Tabla 41. Test de Kruskal-Wallis: errores.

Panel A: <i>Types</i> errores total											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	13,65	12	46,70	,61	6,83	3,00	47,00	221,29	39,420	,000
Bq2	108	11,82	11	42,01	,62	6,48	,00	32,00	192,68		
Bq3	134	8,82	8	30,71	,47	5,54	,00	35,00	140,06		
Panel B: <i>Tokens</i> errores total											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	14,61	13	67,35	,73	8,20	3,00	59,00	219,58	38,329	,000
Bq2	108	12,60	11	49,60	,67	7,04	,00	35,00	194,57		
Bq3	134	9,24	8	37,60	,52	6,13	,00	37,00	140,10		
Panel C: <i>Types</i> errores léxicos											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	10,24	9	26,03	,46	5,10	2,00	31,00	218,38	32,890	,000
Bq2	108	9,02	8	29,03	,51	5,38	,00	25,00	191,06		
Bq3	134	6,91	6	24,03	,42	4,90	,00	31,00	144,03		
Panel D: <i>Tokens</i> errores léxicos											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	10,78	10	34,69	,53	5,89	2,00	38,00	217,87	33,125	,000
Bq2	108	9,39	9	31,90	,54	5,64	,00	27,00	192,23		
Bq3	134	7,10	6	25,39	,46	5,32	,00	33,00	143,55		
Panel E: <i>Types</i> errores funcionales											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	3,39	3	6,60	,23	2,56	,00	15,00	213,33	28,314	,000
Bq2	108	2,79	3	3,60	,18	1,89	,00	9,00	194,00		
Bq3	134	1,94	2	2,77	,14	1,66	,00	8,00	146,29		
Panel F: <i>Tokens</i> errores funcionales											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	3,8374	3	10,957	,29846	3,31013	,00	21,00	210,67	26,643	,000
Bq2	108	3,2037	3	5,341	,22239	2,31112	,00	11,00	196,50		
Bq3	134	2,1418	2	3,626	,16451	1,90430	,00	9,00	146,72		

Tabla 42. Test de Kruskal-Wallis: categoría de palabra léxica.

Panel A: <i>Tokens</i> Sustantivo											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	23,47	22	83,95	,82	9,16	7	57	138,43	36,189	,000
Bq2	108	29,00	27	133,80	1,11	11,56	9	66	192,44		
Bq3	134	31,80	30	168,41	1,12	12,97	5	79	216,29		
Panel B: <i>Tokens</i> adjetivo											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	7,39	7	15,24	,35	3,90	0	27	145,83	23,195	,000
Bq2	108	9,45	9	17,24	,39	4,15	1	20	200,70		
Bq3	134	9,74	9	22,20	,40	4,71	1	25	202,86		
Panel C: <i>Tokens</i> verbo											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	16,86	16	49,59	,63	7,04	5	42	145,07	25,502	,000
Bq2	108	20,21	19	52,58	,69	7,25	5	39	193,07		
Bq3	134	21,84	21,5	78,69	,76	8,87	4	52	209,70		
Panel D: <i>Tokens</i> adverbio											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	7,54	7	15,74	,35	3,96	0	25	151,03	18,543	,000
Bq2	108	9,00	8	15,73	,38	3,96	1	19	190,32		
Bq3	134	9,93	10	23,37	,47	4,83	0	23	206,45		

Tabla 43. Test de Kruskal-Wallis: categoría de palabra léxica.

Panel A: <i>Tokens</i> errores sustantivo											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	3,52	3	6,23	,22	2,49	0	11	203,98	8,423	,015
Bq2	108	3,02	2	6,62	,24	2,57	0	11	179,56		
Bq3	134	2,93	2	9,86	,27	3,14	0	19	166,51		
Panel B: <i>Tokens</i> errores adjetivo											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	1,47	1	2,35	,13	1,53	0	9	192,32	5,824	,05
Bq2	108	1,51	1	2,34	,14	1,53	0	5	193,28		
Bq3	134	1,13	1	1,90	,11	1,38	0	7	166,16		
Panel C: <i>Tokens</i> errores verbo											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	4,81	5	10,84	,29	3,29	0	20	222,34	41,831	,000
Bq2	108	4,03	3,5	9,96	,30	3,15	0	15	192,69		
Bq3	134	2,47	2,4	4,43	,18	2,10	0	9	139,09		
Panel D: <i>Tokens</i> errores adverbio											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	1	1	1,47	,10	1,21	0	6	200,44	9,784	,008
Bq2	108	,87	0	1,25	,10	1,11	0	4	187,48		
Bq3	134	,56	0	,66	,07	,81	0	3	163,38		

Tabla 44. Test de Kruskal-Wallis: tipo del error.

Panel A: Variable errores formales											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	8,21	7	25,82	,45	5,08	1	33	218,63	30,813	,000
Bq2	108	6,71	7	19,98	,43	4,47	0	21	188,16		
Bq3	134	5,16	4	18,87	,37	4,34	0	26	146,14		
Panel B: Variable errores semánticos											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	1,21	1	1,36	,10	1,16	0	5	185,09	,723	,697
Bq2	108	1,47	1	3,61	,18	1,90	0	12	187,75		
Bq3	134	1,18	1	1,74	,11	1,32	0	5	177,25		
Panel C: Variable errores mixtos											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	1,38	1	2,12	,13	1,45	0	6	204,70	18,663	,000
Bq2	108	1,26	1	2,19	,14	1,48	0	7	194,44		
Bq3	134	,75	0	1,27	,09	1,12	0	5	153,87		

Tabla 45. Test de Kruskal-Wallis: origen del error.

Panel A: Errores Interlingüísticos											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	5,94	5	20,87	,41	4,56	0	32	221,54	28,692	,000
Bq2	108	4,12	3	10,17	,30	3,18	0	15	177,82		
Bq3	134	3,48	3	11,66	,29	3,41	0	19	151,79		
Panel B: Errores Intralingüísticos											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	4,34	4	5,88	,21	2,42	0	11	198,77	17,465	,000
Bq2	108	4,74	4	11,20	,32	3,34	0	17	202,32		
Bq3	134	3,29	3	6,31	,21	2,51	0	14	152,96		
Panel C: Errores Ambiguos											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	,52	0	,56	,06	,75	0	3	192,40	8,594	,014
Bq2	108	,58	0	,84	,08	,91	0	5	194,14		
Bq3	134	,32	0	,50	,06	,71	0	5	165,39		



Tabla 46. Test de Kruskal-Wallis: error interlingüístico según el tipo.

Panel A: Interlingüístico/Forma											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	4,43	3	14,85	,34	3,85	0	27	222,95	30,243	0,000
Bq2	108	2,76	2	6,12	,23	2,47	0	12	175,88		
Bq3	134	2,39	2	7,41	,23	2,72	0	15	152,07		
Panel B: Interlingüístico/Contenido											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	,71	0	,99	,08	,99	0	5	187,26	0,596	0,742
Bq2	108	,66	0	1,29	,10	1,13	0	6	177,74		
Bq3	134	,67	0	,97	,08	,98	0	5	183,34		
Panel C: Interlingüístico/Mixto											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	,78	0	1,250	,10081	1,11805	0	5	198,62	10,804	0,005
Bq2	108	,69	0	1,130	,10229	1,06304	0	5	190,15		
Bq3	134	,41	0	,696	,07208	,83440	0	4	162,9		

Tabla 47. Test de Kruskal-Wallis: error intralingüístico según el tipo.

Panel A: Intralingüístico/Forma											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	3,52	3	5,21	,20	2,28	0	11	200,06	13,686	0,001
Bq2	108	3,67	3	8,67	,28	2,94	0	16	196,47		
Bq3	134	2,58	2	4,83	,18	2,19	0	12	156,49		
Panel B: Intralingüístico/Contenido											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	,47	0	,48	,06	,69	0	3	178,21	3,305	0,192
Bq2	108	,76	0	1,35	,11	1,16	0	6	196,43		
Bq3	134	,47	0	,50	,06	,71	0	3	176,57		
Panel C: Intralingüístico/Mixto											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	,33	0	,42	,05	,64	0	3	189,11	3,193	0,203
Bq2	108	,29	0	,32	,05	,56	0	3	187,73		
Bq3	134	,22	0	,28	,04	,52	0	2	173,57		

Tabla 48. Test de Kruskal-Wallis: error ambiguo según el tipo.

Panel A: Ambiguo/Forma											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	,24	0	,26	,04	,51	0	2	187,07	3,182	0,204
Bq2	108	,26	0	,32	,05	,57	0	3	189,02		
Bq3	134	,17	0	,32	,04	,57	0	5	174,41		
Panel B: Ambiguo/Contenido											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	,0163	,0000	,01	,01	,12	0	1	180,47	1,219	0,544
Bq2	108	,0370	,0000	,03	,01	,18	0	1	184,26		
Bq3	134	,0373	,0000	,03	,01	,19	0	1	184,31		
Panel C: Ambiguo/Mixto											
Bloques	N	Media	Mediana	Varianza	Error estándar	Desviación estándar	Mínimo	Máximo	Rango medio	Chi-cuadrada	Valor-p
Bq1	123	,2602	0	,22	,04	,47	0	2	193,57	9,217	0,010
Bq2	108	,2778	0	,37	,05	,60	0	3	189,09		
Bq3	134	,1119	0	,11	,02	,33	0	2	168,39		

## Anexo 8

*Corpus* lingüístico en formato digital (CD).