



DIVERSIDAD LÉXICA EN EL DISCURSO ORAL DE ESTUDIANTES DE ESPAÑOL COMO LENGUA EXTRANJERA (ELE) DE NIVEL C1

*Lexical diversity in Spanish as
a second language students' speech (level C1)*

Patricia Guillén Solano¹

RESUMEN

En este artículo se analiza la diversidad léxica en el discurso oral de estudiantes de español como lengua extranjera de nivel C1. Luego de la aplicación de la fórmula Pd50, se calcula el índice de diversidad léxica en cinco presentaciones orales realizadas en el ámbito universitario. Con base en los resultados, y considerando además algunos factores de índole cualitativo, la diversidad léxica se analiza tomando en cuenta los descriptores de competencia léxica establecidos por el *Marco Común Europeo de Referencia para las Lenguas* (2002) y por el *Plan Curricular del Instituto Cervantes* (2006). A partir de este análisis, se concluye que la diversidad léxica de las presentaciones orales no cumple con dichos descriptores, a la vez que se proponen posibles razones de esta inconsistencia. Finalmente, se discute la relación entre diversidad léxica y falta de competencia en los géneros discursivos en el ámbito educativo.

Palabras clave: Diversidad léxica, discurso expositivo, español como lengua extranjera, competencia léxica, ámbito educativo.

ABSTRACT

In this article lexical diversity in Spanish as a foreign language spoken discourse of C1 level students is analyzed. By applying the formula Pd50, the index of lexical diversity is calculated on the basis of five oral presentations within a university setting. According to the results, and considering some complementary qualitative factors, lexical diversity is analyzed taking into account the descriptors for lexical competence established by the *Marco Común Europeo de Referencia para las lenguas* (2002) and the *Plan Curricular del Instituto Cervantes* (2006). From this analysis it is concluded that the lexical diversity of the oral presentations does not meet those descriptors, while possible reasons for this inconsistency are proposed. Finally, the relationship between lexical diversity and lack of competence in discourse genres in educational settings is discussed.

Key Words: Lexical diversity, expository discourse, Spanish as a foreign language, lexical competence, educational setting.

¹ Universidad de Costa Rica, Departamento de Lingüística. Costa Rica.

Correo electrónico: patricia.guillen@ucr.ac.cr

Recepción: 05-05-2019. Aceptación: 26-03-2020.



1. Introducción

Tal como apuntan Bellés y Fortanet (2005), la internacionalización de la educación superior ha derivado en los últimos años en una preocupación constante por desarrollar más investigación sobre las características del discurso oral que desarrollan los estudiantes de español como lengua extranjera (en adelante, ELE) inmersos en el ámbito universitario.

En relación con estas características, parece claro que los estudiantes consideran el uso de un léxico variado y preciso, acorde con la situación comunicativa, una parte fundamental de su competencia en la lengua meta. Así lo evidencia el estudio de Gil del Moral (2017) sobre el aprendizaje del español académico. La autora determina, a través de un cuestionario de necesidades del alumnado, el alto grado de interés de los estudiantes por la enseñanza explícita de estrategias que ayuden a mejorar la competencia léxica, es decir, el conocimiento del vocabulario y la capacidad para utilizarlo adecuadamente (Instituto Cervantes, 2002, p. 208).

En contraste con la abundancia de estudios sobre la riqueza léxica en español como lengua materna (Ávila, 1986, 1988; Morales, 1986; Haché, 1991; López Morales, 1995-1996; Barahona Novoa, 1996; Leoni de León, 1997; Murillo Rojas, 1998, 2009; Araya, 2012; Madrigal y Vargas, 2016, entre muchos otros), las investigaciones en español como lengua extranjera son más escasas, y muchas de ellas se centran en el discurso escrito. Berton (2014), por ejemplo, estudia la diversidad léxica de la producción escrita en español de jóvenes adultos suecos en un contexto de adquisición formal (grado de bachillerato), en niveles bajos, intermedios y avanzados de dominio de la lengua. Comparando la diversidad léxica en los niveles A1-A2 y B1-B2, determina que las producciones escritas de los estudiantes de nivel más alto presentan una diversidad léxica mayor que la de los estudiantes de niveles más bajos, y que, por lo tanto, el nivel de competencia en la lengua meta ejerce un efecto sobre la diversidad léxica. Aun cuando este estudio es de carácter comparativo entre niveles



y se centra en el discurso escrito, mientras que la presente investigación se basa en un solo nivel (C1) y en el discurso oral, la sistematización de conceptos que plantea el autor resulta de gran utilidad.

También en la línea de la producción escrita, Vargas Castro (2015) analiza el uso del léxico en redacciones de estudiantes de ELE de nivel B2 a partir de cuatro parámetros: variedad, propiedad, precisión y competencia. La autora concluye que la mayoría de deficiencias se relacionan con la variedad léxica. Esta conclusión se retomará más adelante, pues resulta de particular importancia para el presente estudio.

En cuanto al discurso oral, destaca el estudio de Barkelius (2007), quien compara la riqueza léxica de hablantes nativos chilenos con la de hablantes suecos de ELE de nivel avanzado, todos con educación superior universitaria, utilizando para ello un corpus obtenido a partir de la narración de una historia, página por página, en un libro de imágenes. Mediante el cálculo de la razón entre tipos y lemas, determina que los promedios son similares entre hablantes nativos y no nativos (0,84). Si bien la presente investigación no es de tipo comparativo entre L1 y L2, los resultados de este estudio son de especial interés por cuanto los dos grupos de informantes muestran promedios muy semejantes en relación con el índice de diversidad léxica, de manera que los hablantes no nativos se sitúan al nivel de los nativos.

Partiendo entonces de que el estado de la investigación sobre la diversidad léxica en ELE evidencia la escasez de estudios centrados en el discurso oral, y de que el dominio de un léxico variado se considera fundamental para el desarrollo de este tipo de discurso, el objetivo de este trabajo es determinar la diversidad léxica de cinco presentaciones orales que desarrollan cinco temas distintos: el cambio climático (P1), la pobreza (P2), la relación entre el poder económico y el Estado (P3), la noción de tiempo como diferencia cultural (P4) y el impacto de la televisión en la sociedad actual (P5).

La estimación de la diversidad léxica se lleva a cabo con base en la diversidad de palabras de contenido (sustantivos, adjetivos y verbos) en presentaciones orales, utilizando la fórmula Pd50 –



promedio de palabras diferentes—: conteo de las palabras diferentes en textos de cincuenta enunciados, que se calcula dividiendo el número total de vocablos tipo por el número total de palabras (Murillo Rojas, 2009, p. 130). Se consideran, a su vez, otros factores de índole cualitativo con el fin de determinar si el discurso oral de hablantes de ELE inmersos en el ámbito universitario cumple con los descriptores de competencia léxica establecidos por el *Marco Común Europeo de Referencia para las Lenguas* (2002) y por el *Plan Curricular del Instituto Cervantes* (2006). Se parte de tres premisas:

- El nivel de competencia en la lengua meta ejerce un efecto sobre la diversidad léxica.
- En el caso del nivel C1, tal como ocurre en el discurso nativo, la mayoría de deficiencias relacionadas con la competencia léxica se relacionan con la diversidad y, en menor medida, con la precisión y la propiedad léxicas.
- Existe una relación entre diversidad léxica y grado de competencia en un género discursivo particular.

2. Nociones preliminares

Antes de iniciar con el análisis de las muestras de discurso oral, se considera importante precisar cierta terminología que resulta clave para el estudio. Se abordará, en primer lugar, los términos *léxico* y *vocabulario*, utilizados ampliamente en el ámbito de la enseñanza de lenguas extranjeras.

En los estudios lexicográficos especializados, se suele establecer una distinción entre léxico y vocabulario: el léxico se refiere al conjunto de palabras que un hablante sabe en un momento determinado, por lo que resulta un término más global, mientras que el vocabulario alude a los vocablos utilizados en un acto de habla concreto (Justicia, 1995, p. 11).

En la didáctica de la lengua, y específicamente en el ámbito de la enseñanza de ELE, no obstante, los términos vocabulario y léxico se consideran equivalentes, en cuanto corresponden al



“conjunto de unidades léxicas de una lengua”, las cuales pueden ser simples –una palabra– o complejas –dos o más palabras con sentido unitario– (Centro Virtual Cervantes, 2020).

Desde un punto de vista práctico, las unidades léxicas se convierten en las unidades base en el aprendizaje de vocabulario e implican la unión de, al menos, una forma léxica con un valor semántico que pueda ser identificado (Higueras García, 1997; Izquierdo, 2005). La adquisición de las unidades léxicas se asocia, a su vez, con una serie de subaprendizajes (Izquierdo, 2005), entre ellos: el reconocimiento de los distintos significados de cada unidad y la capacidad de producirla según el contexto, el reconocimiento y la producción de otras unidades con las que se relacione desde el punto de vista del significado (sinónimos, antónimos, hiperónimos), así como su adecuación pragmática, es decir, la capacidad de adaptarla considerando el lugar en el cual se produce el intercambio, la intención de este y el interlocutor al que se dirige. Estos subaprendizajes implican, de una u otra forma, la noción de diversidad léxica.

En cuanto al término *diversidad léxica*, Carroll (1964) lo define como “la cantidad relativa de repetición de vocabulario” (p. 379). En esta misma línea, Durán, Malvern, Richards y Chipere (2004) se refieren a la diversidad léxica como “*the range of vocabulary and avoidance of repetition*” (p. 3). Para Jarvis (2013), esta consiste en “*the variety of word use*” (p. 100), de ahí que los términos diversidad léxica y variedad léxica se utilicen como sinónimos.² En la presente investigación se analiza la diversidad léxica en el discurso oral en el ámbito universitario de estudiantes de ELE, tomando como base la diversidad en el uso de palabras de contenido en presentaciones orales. Al respecto, conviene realizar algunas precisiones:

² Aun cuando varios autores (Laufer y Nation, 1995; Yu, 2010; Jarvis, 2013) utilicen los términos *riqueza léxica* y *diversidad léxica* como equivalentes, en este trabajo se sigue la perspectiva de Read (2000, p. 200) en cuanto a que en la riqueza léxica convergen la diversidad léxica, la sofisticación léxica y la densidad léxica, entre otros posibles constructos.



- En cuanto a la variedad de discurso oral, se debe puntualizar que se trata de presentaciones orales desarrolladas en un curso de expresión oral dirigido a estudiantes de ELE de nivel C1. Estas presentaciones corresponden a un discurso monológico expositivo, el cual se caracteriza por informar y transmitir datos de forma organizada y jerarquizada e incluye, necesariamente, la explicación –hacer comprender por qué un fenómeno es o actúa de un modo determinado (Bassols y Torrent, 2012)– y la argumentación –hacer que un auditorio comparta opiniones relativas a un tema determinado– (Adam, 1992).
- En relación con el ámbito en el cual se ubica el discurso, es importante señalar que, de acuerdo con el Marco Común Europeo de Referencia para las Lenguas (Instituto Cervantes, 2002, p. 15), las actividades de la lengua se encuentran contextualizadas dentro de cuatro ámbitos generales: el público, el personal, el educativo y el profesional. En este sentido, las muestras seleccionadas se ubican en el ámbito educativo, el cual comprende el contexto de aprendizaje e implica la adquisición de conocimientos o destrezas específicas que, en el caso particular de esta investigación, están relacionados con el uso variado del léxico en el discurso oral de tipo expositivo.
- Por último, se debe aclarar que la diversidad léxica se determinará con base en el léxico disponible, es decir, el léxico correspondiente a una situación comunicativa concreta, lo que implica que esté basado en un tema específico.³ En este caso, dicho tema fue propuesto por el estudiante encargado de desarrollar la presentación oral.

³ El término léxico disponible se opone al de léxico básico: Mientras el léxico básico abarca los vocablos más usuales de una comunidad y se caracteriza por un alto grado de estabilidad –permanencia– que le permite aparecer con mucha frecuencia y en todo tipo de discursos, independientemente de la temática de que se trate, el léxico disponible estaría integrado por aquellas palabras que acuden a la mente del hablante como reacción a un estímulo temático concreto (López Chávez, 1994, p. 69).



3. Metodología

Para llevar a cabo el análisis detallado de las muestras y determinar su diversidad léxica, se seleccionaron cinco grabaciones en audio que corresponden a cinco presentaciones orales realizadas por estudiantes de ELE en un curso de expresión oral impartido en el Programa de Español como Lengua Extranjera de la Universidad de Costa Rica. Las grabaciones estuvieron a cargo de la investigadora y del señor Luis Serrato Pineda, profesor de este programa.

Las cinco grabaciones fueron seleccionadas tomando en consideración los siguientes criterios:

- Ninguno de los cinco expositores se refiere a soportes visuales. Este criterio resulta importante dado que se trata de grabaciones en audio y, de no ser así, se carecería de los elementos necesarios para valorar el discurso de manera global.
- Los cinco expositores poseen un nivel C1 de acuerdo con los parámetros generales del MCER. Este nivel fue determinado por tres profesores de ELE, todos costarricenses hablantes nativos de español con una amplia experiencia en docencia y en la realización de exámenes de ubicación, graduados de filología española y con estudios de maestría en el área de segundas lenguas y metodologías de enseñanza, siguiendo los criterios establecidos por el *Marco Común Europeo de Referencia para las Lenguas* (2002, pp. 75-78) para la interacción oral y la conversación formal, que constituyen las habilidades por examinar:
 - C1: Se expresa con fluidez y espontaneidad, casi sin esfuerzo. Domina un amplio repertorio léxico que le permite suplir sus deficiencias fácilmente con circunloquios. Apenas hay una búsqueda evidente de expresiones o estrategias de evitación; solo un tema conceptualmente difícil puede obstaculizar un discurso fluido y natural.
 - Comprende con facilidad las interacciones complejas que tratan temas abstractos, complejos y desconocidos y que tienen lugar entre terceras partes en grupos de discusión. Es perfectamente capaz de participar en ellas.



- Sigue el ritmo de un debate con facilidad, incluso sobre temas abstractos, complejos y desconocidos. Argumenta su postura formalmente y con convicción, respondiendo a preguntas y comentarios y contestando de forma fluida, espontánea y adecuada a argumentaciones complejas contrarias.
- Las grabaciones se llevaron a cabo con consentimiento informado previo de los profesores titulares del curso y de los estudiantes, por lo tanto, sabían que estaban siendo grabados. Con el fin de disminuir los posibles factores relacionados con este hecho (mayor nerviosismo, menor naturalidad), se comenzó a grabar minutos antes del inicio de cada presentación, y se colocó la grabadora fuera de la vista de los participantes.
- Cada grupo estaba conformado por entre cuatro y siete estudiantes. El número de estudiantes por grupo no corresponde a una selección de la investigadora, sino a una decisión del Programa de Español como Lengua Extranjera de la Universidad de Costa Rica, que conforma los grupos de acuerdo con el nivel de los participantes.

El corpus recogido corresponde a un discurso propio del ámbito educativo universitario, que conlleva al uso de una norma culta, pero sin la especialización léxica de una carrera universitaria o de una profesión en específico. Todos los participantes cuentan con estudios universitarios y se encuentran en un programa de inmersión lingüística en lengua española en Costa Rica, su nivel de dominio de la lengua es C1 de acuerdo con los parámetros establecidos por el MCER y todos tienen como lengua materna el inglés.

En la Tabla 1, se detalla la cantidad total de palabras⁴ y de minutos que conforman el corpus grabado y transcrito.

⁴ El término *palabra* se utiliza en el presente estudio con el sentido que se le asigna en el ámbito de análisis léxico estadístico: unidad gráfica separada de los grupos vecinos por un blanco o una puntuación (Müller, 1973, p. 226). Esta unidad consta de un número entero de sílabas y sus límites pueden estar definidos por la función demarcativa de algunos de sus fonemas (Lara, 2006, p. 138).

**Tabla 1.** Detalle del corpus en estudio

Presentaciones	Número de palabras	Minutos	Palabras por minuto
Presentación 1	613	7,5	82
Presentación 2	602	7	86
Presentación 3	611	8	76
Presentación 4	575	6,5	88
Presentación 5	682	5	136
Total de palabras y minutos registrados	3083	34	

Fuente: Elaboración propia.

En el análisis se toman en cuenta las variables que se presentan a continuación.

3.1. Variables

3.1.1. Diversidad léxica

Se entenderá como diversidad léxica la diversidad en el uso de palabras de contenido por parte de los hablantes de ELE, en producciones orales de 575 a 682 palabras.

Tomando en cuenta que Durán *et al.* (2004) advierten sobre la dificultad de alcanzar conclusiones válidas con respecto a la variedad léxica cuando se comparan textos de extensión diferente, pues cuantas más palabras haya en el texto (*lemas*), más posibilidades hay de que sean repetidas (menos *tipos*), se ha tratado de que la extensión de las muestras sea similar. Para llevar a cabo el análisis, se han tomado las muestras contemplando los cincuenta primeros enunciados de cada presentación, con el fin de aplicar la fórmula Pd50. Esta fórmula es utilizada por Murillo Rojas (2009) para medir la diversidad léxica en narraciones orales de niños, y se considera que se adapta muy bien al objetivo de investigación, dado que resulta más conveniente segmentar las presentaciones orales en enunciados que limitarlas forzosamente a cincuenta o cien palabras, tal como plantean otros métodos de estandarización (López Morales, 2010; Ávila, 1986).



3.1.2. Género discursivo

Los discursos se inscriben en géneros identificables según una serie de pautas y de convenciones que los hablantes siguen de acuerdo con la práctica social en la que participan (Calsamiglia y Tusón, 2012). En el caso del presente trabajo, el género corresponde a la presentación oral, una de las manifestaciones del discurso que comparte las características fundamentales de la lengua oral en el ámbito educativo: se construye a partir de un tema especializado y, por tanto, es planificada, presenta un tono formal, es más informativa que interactiva, su canal es oral-auditivo y es simultánea en el espacio y en el tiempo (Castellà y Vilà, 2005). El tema de cada presentación oral fue propuesto por el estudiante expositor, de manera que se desarrollaron cinco temáticas: el cambio climático (P1), la pobreza (P2), la relación entre el poder económico y el Estado (P3), la noción de tiempo como diferencia cultural (P4) y el impacto de la televisión en la sociedad actual (P5).

3.1.3. Edad

Considerando que al trabajar con una población de participantes de edades diferentes, el número total de palabras suele aumentar conjuntamente con la edad, tanto en la producción escrita como en la oral (Vermeer, 2000; Johansson, 2008), se ha limitado el estudio a participantes de veintitrés a veintiséis años, pues no se contaba con suficientes sujetos para establecer una comparación basada en esta variable.

3.1.4. Nivel de estudios

En vista de que se trata de determinar la diversidad léxica en el discurso oral de estudiantes de ELE inmersos en el ámbito educativo, se estableció que todos los participantes contaran con al menos un año de estudios en ELE en el nivel universitario.



3.1.5. Lengua materna

Todos los participantes tienen como lengua materna el inglés. Dado el tamaño de la muestra, se decidió que todos tuvieran la misma lengua materna para minimizar diferencias que pueden darse debido a esta variable, pues su análisis está fuera de los objetivos de esta investigación.

3.2. Procedimientos metodológicos y análisis del corpus

En primer lugar, se transcribieron las grabaciones en ortografía estándar utilizando los signos del sistema de transcripción del Grupo Val.Es.Co., dirigido por Briz (2000). Para llevar a cabo la transcripción de las grabaciones, se utilizó el programa digital *Express Scribe*.

Una vez transcrito el corpus, se seleccionaron los cincuenta primeros enunciados de cada presentación. El análisis detallado del corpus se realizó mediante el programa *MAXQDA*, un *software* para el análisis cualitativo de datos que ofrece herramientas básicas para la organización y el análisis de textos, las cuales nos permitieron codificar el léxico a partir de los siguientes criterios:

1. Palabras de contenido: Siguiendo a Murillo Rojas (2009), únicamente se contabilizaron como palabras de contenido sustantivos, adjetivos y verbos. Por lo tanto, no se tomaron en cuenta preposiciones, conjunciones, pronombres, artículos ni adverbios. Si bien se trata de estudiantes de ELE y tanto las palabras de contenido como las gramaticales forman parte de la competencia léxica de acuerdo con el MCER (Instituto Cervantes, 2002, p. 108), se valora la diversidad léxica, no la competencia léxica general.
2. Se excluyeron del conteo las palabras truncadas, propias de los falsos comienzos y las pausas rellenas (muletillas).
3. Para obtener el número de verbos distintos, se contaron una sola vez las formas infinitivas de cada expresión verbal. Las perífrasis verbales se conciben como una unidad verbal y, por lo tanto, cada una se contabiliza como un solo verbo.



4. En el caso de los sustantivos y los adjetivos, se contaron los lexemas una sola vez. No se consideraron los nombres propios.
5. Una vez contabilizadas las palabras de contenido, se aplicó la fórmula Pd50.

4. Resultados

4.1. Índice de diversidad léxica

En la Tabla 2, se exponen los resultados:

Tabla 2. Índice de diversidad léxica

Presentaciones	Número de palabras de contenido	Número de palabras distintas (tipo)	Índice
Presentación 1	183	91	0,50
Presentación 2	266	144	0,54
Presentación 3	238	91	0,38
Presentación 4	239	95	0,40
Presentación 5	253	133	0,53
Media	235,8	110,8	0,47

Fuente: Elaboración propia.

Siguiendo los parámetros que establecen que los índices más cercanos a 1 se asocian con una diversidad léxica alta, así como los criterios señalados por Murillo Rojas (2009): mayor a 0,75, superior; 0,56 a 0,74, normal; menor a 0,55, inferior, en el caso de este corpus los índices muestran una variedad léxica inferior, pues el índice más alto es de 0,54, con una media de 0,47.

Este índice se aleja de los resultados obtenidos en los estudios sobre riqueza léxica realizados por Barahona Novoa (1996), con una media de 67,07 en textos producidos en radio, y Leoni de León (1997), con una media de 68,69 para la prensa escrita. Los resultados del presente estudio están más cerca de los obtenidos por Madrigal y Vargas (2016), con una media de 56,97 en textos escritos producidos por estudiantes universitarios. Habría que tomar en cuenta, eso sí, que el cálculo de la



riqueza léxica en los tres estudios anteriores se realizó con una metodología distinta a la que se ha aplicado en esta investigación. En el caso de Murillo Rojas (2009), quien aplica la fórmula Pd50, la media es de 0,65 para textos producidos por niños prescolares.

Por lo tanto, comparativamente la diversidad léxica del discurso desarrollado en las presentaciones orales de estudiantes de ELE inmersos en el ámbito universitario es baja, al no superar en ningún caso el índice de 0,54, con una media de 0,47, significativamente inferior a la media establecida por los investigadores citados.

4.2. Algunas consideraciones de índole cualitativo

Varios autores (Berton, 2014; Jarvis, 2013; Higuera García, 1997) señalan la necesidad de complementar el análisis numérico de la diversidad léxica con un análisis cualitativo de diversos factores que afectan el uso del léxico en el discurso. A continuación, se da cuenta de algunos de ellos.

4.2.1. Repeticiones

En vista de que, en términos generales, la diversidad léxica se concibe como falta de repetición, se considera importante realizar una serie de apuntes a partir de un análisis cualitativo de los datos:

Los hablantes utilizan cinco tipos de repeticiones de palabras (en este caso se consideraron también las repeticiones de palabras funcionales o gramaticales):

- Repeticiones que se deben a la falta de variedad léxica, en el sentido de que se repite excesivamente una misma palabra y no se recurre al uso de unidades léxicas con contigüidad semántica, ni tampoco a la paráfrasis, a la definición o al circunloquio con el fin de evitar la repetición:
 - P1: *el **altitud** aquí en San José / es/ mm/ más alto que en mi estado/ mm/ estamos en la **altitud**// cerca de mil doscientos metros/ pero/ mm/ en South Dakota/ *el **altitud** es más bajo como/ mm/ quinientos metros/



- Repeticiones que tienen como único propósito evitar un vacío en el discurso, es decir, que están relacionadas con la fluidez:
 - P1: **las las** diferencias/ mm// que existen//
- Repeticiones que se utilizan como estrategia discursiva de intensificación:
 - P1: donde yo vivo es **muy muy** plano no hay montañas
- Repeticiones que se utilizan como estrategia discursiva de recuperación del referente con fines argumentativos y explicativos:
 - P1: En North Dakota y South Dakota/ hay/ **industrias** (...) Estas **industrias**
- Repeticiones que constituyen una estrategia de autocorrección:
 - P1: **muchas**/ mm/ turistas- **muchos** turistas/

En la Tabla 3, se cuantifican las repeticiones de acuerdo con las especificaciones anteriores:

Tabla 3. Repeticiones de acuerdo con su función en el discurso

Presentaciones	Evitación de vacío ante falta de unidad léxica alternativa	Evitación de vacío ante falta de fluidez en el discurso	Intensificación	Recuperación del antecedente	Autocorrección
Presentación 1	1	8	2	2	6
Presentación 2	4	11	0	2	5
Presentación 3	1	8	5	2	6
Presentación 4	5	11	2	1	0
Presentación 5	0	25	4	3	0
TOTAL	11	63	13	10	17

Fuente: Elaboración propia.

Como se evidencia en los datos, el tipo de repetición más frecuente en el corpus se produce con el fin de evitar un vacío en el discurso, de manera que cumple la misma función de las muletillas en el sentido que plantea Vigara (2005, p. 28): se utiliza para evitar el silencio o la desconexión



abrupta entre los enunciados, y su uso es abundante, inconsciente e irreflexivo. Este tipo de repetición, junto con la que responde, en principio, a una carencia en el área léxica, representa el 65 % de los casos presentes en el corpus, mientras que el 35 % de las repeticiones corresponde a estrategias discursivas de autocorrección, de intensificación o de recuperación del referente.

Cabe señalar que sobre este punto se debe considerar la situación comunicativa en la que se produce el discurso: se trata de una presentación oral dentro de un salón de clase, situación que en sí misma implica nerviosismo y, por lo tanto, titubeos y autocorrección. El hecho de que la finalidad del acto comunicativo parta del cumplimiento de una tarea académica hace que se produzcan repeticiones de diferente tipo, propias también del discurso de hablantes nativos sometidos a las mismas condiciones (v. Camacho Adarve, 2009). Adicionalmente, en el caso de los hablantes de ELE, al conocimiento de la temática que se aborda y a la adecuada organización del discurso debe sumársele el nivel de dominio de los distintos componentes de la lengua: morfosintáctico, fonético, léxico, semántico y pragmático, lo cual aumenta la probabilidad de que se recurra a la repetición de forma más frecuente.

4.2.2. La precisión y su relación con la diversidad léxica

La precisión suele definirse según el objetivo que persigue: evitar el uso de vocablos con significados muy generales a cambio de vocablos con significados suficientemente específicos. Se ha considerado el uso de las denominadas *palabras comodines* o *proformas léxicas* como un aspecto relacionado con la precisión léxica que puede afectar el índice de diversidad de palabras de contenido, tal como se evidencia en el siguiente ejemplo:

- P1: con/// mm// escogí// eh/ el tema de cambio climático porque// mm// ess// un tema que me interesa mucho/ y también me asusta// mm/ y también/ ess/ una **cosa** que/ quiero/ estudiar/ en el futuro/

En la Tabla 4, se indica el número de apariciones de cada proforma para cada una de las cinco presentaciones.



Tabla 4. Uso de proformas léxicas

Proformas	‘cosa’	‘cosas’	‘eso’	‘algo’	‘esto’
Presentación 1	1	0	0	0	0
Presentación 2	1	0	3	0	0
Presentación 3	0	2	1	1	1
Presentación 4	2	3	0	2	0
Presentación 5	2	1	1	0	0
TOTAL	6	6	5	3	1

Fuente: Elaboración propia.

Como se constata en los datos expuestos, el uso de palabras comodín o proformas léxicas es bajo, pues no sobrepasa las seis apariciones por presentación. En términos globales, representa tan solo el 0,68 % del total de palabras del corpus.

4.2.3. La propiedad y su relación con la diversidad léxica

La propiedad se relaciona con la selección adecuada de los vocablos según la intención discursiva que se persigue, de manera que la palabra seleccionada posea una acepción próxima al significado que se quiere alcanzar. Se considera relevante dar cuenta de los casos relacionados con la propiedad léxica que aparecen en el corpus:

- Cuatro palabras utilizadas en la lengua materna de los hablantes (en este caso, el inglés): ‘and’ y ‘not’ en la primera presentación, y ‘unions’, ‘air’, ‘traffic’ y ‘controlers’, en la tercera presentación.
- Dos palabras utilizadas con un valor distinto al que tienen en la lengua meta:
 - ‘Valer’ en lugar de ‘valorar’:
 - P1: la gente aquí en/ en Costa Rica/ **vale**/ mm/ la naturaleza
 - ‘Provenir’ en lugar de ‘proveer’:
 - P2: gobiernos pueden/ **provenir** educación/ servicios/

Se producen también seis peticiones de ayuda al profesor, sea mediante una pregunta directa o una petición indirecta, y que tienen como objetivo, precisamente, cerciorarse de que la palabra seleccionada se ajusta al significado que se quiere expresar, como ocurre en el siguiente ejemplo:



- P1: Estudiante: / ¿petro?/ no sé §

Prof: § petróleo

5. Valoración de la diversidad léxica a la luz de los descriptores del Marco Común Europeo de Referencia para las Lenguas (MCER) y del Plan Curricular del Instituto Cervantes (PCIC)

Aun cuando se ha comprobado que el nivel de competencia en la lengua meta ejerce un efecto sobre la diversidad léxica (Berton, 2014), cabría preguntarse si tanto los estudiantes universitarios que tienen como lengua materna el español como aquellos para quienes corresponde a una lengua extranjera con un nivel de dominio C1 cumplen con el criterio establecido por el Plan Curricular del Instituto Cervantes para dicho nivel (Instituto Cervantes, 2002): “Los estudiantes cuentan con un repertorio léxico extenso y preciso que les permite expresar matices de significado eliminando ambigüedades y confusiones” (p. 20).

Durante el desarrollo de las presentaciones orales, el uso del léxico no condujo a ambigüedades ni a confusiones, sin embargo, con una media de 0,47, y contemplando los otros aspectos considerados para valorar la diversidad léxica, los resultados muestran que el léxico utilizado no resulta lo suficientemente variado para una presentación oral de estudiantes de nivel C1. Se cree, siguiendo a Durán *et al.* (2004), que en los niveles avanzados de aprendizaje de una lengua el grado de precisión léxica de los estudiantes podría ser mayor, lo que explicaría los índices bajos de diversidad. No obstante, se han identificado una serie de fallos relativos a la precisión léxica. Por lo tanto, tampoco este criterio se cumple satisfactoriamente.

El MCER, por su parte, incluye los siguientes descriptores para valorar la competencia léxica de un estudiante de nivel C1 (Instituto Cervantes, 2002):



Riqueza de vocabulario:

C1: Tiene un buen dominio de un amplio repertorio léxico que le permite superar con soltura sus deficiencias mediante circunloquios; apenas se le nota que busca expresiones o que utiliza estrategias de evitación. Buen dominio de expresiones idiomáticas y coloquiales.

Dominio del vocabulario:

C1: Pequeños y esporádicos deslices, pero sin errores importantes de vocabulario (p. 109).

En este punto, se observa nuevamente el criterio de la amplitud del repertorio léxico, el cual, como ya se ha comentado, no se cumple satisfactoriamente. Laufer (1991, p. 445) propone una posible explicación para la falta de variedad léxica en el discurso de hablantes que se encuentran en niveles altos de manejo de un idioma: al tener un dominio relativamente avanzado de la lengua, los estudiantes no consideran como imperiosa la necesidad de expandir su vocabulario. Son conscientes de que su repertorio léxico les es suficiente para desenvolverse con éxito en un rango amplio de situaciones, así que no se esfuerzan en ampliarlo. Posiblemente por esta razón, el descriptor relacionado con el dominio de vocabulario sí se cumple satisfactoriamente: no se producen en el corpus errores importantes de vocabulario, únicamente hay muestras aisladas de “pequeños y esporádicos deslices”. El problema se presenta en la amplitud del repertorio léxico: el estudiante comete pocos errores precisamente porque no se arriesga a ampliar su repertorio.

La explicación anterior resulta valiosa por cuanto recuerda que las implicaciones que tiene el dominio funcional de la lengua deben ser consideradas cuando se valora la diversidad léxica del discurso oral de estudiantes de ELE inmersos en el ámbito universitario. Según Villar (2011), se debe tener en cuenta que el discurso se sitúa dentro de un *continuum* en cuyos extremos aparecerían géneros prototípicos escritos y orales, en este caso, en un extremo la escritura académica y, en el otro, la conversación coloquial, cuyo índice de riqueza léxica, por cierto, no suele alcanzar el 0,1 (Terrádez, 2000). Vilà (2005), precisamente, denomina *géneros intermedios* a los discursos propios del ámbito de las clases universitarias y que, por lo tanto, se vinculan con la comunicación formal.



Este vínculo con la comunicación formal implica que la construcción y la ejecución del discurso de las presentaciones orales requieran de una formación específica que implica una intervención didáctica en la que se promuevan estrategias para enriquecer el léxico del discurso oral en el aula, partiendo de las características del género discursivo y de la situación de comunicación. Sobre este punto, conviene recordar que Ávila (1988) determina que el factor sociológico que tiene mayor influencia sobre la densidad léxica es el nivel cultural, por encima de factores como la edad o el sexo.

6. Discusión final y conclusiones

El estudio realizado constituye un aporte al campo de la investigación sobre la diversidad léxica en el discurso oral en ELE en el ámbito educativo, campo que, como se ha establecido, no ha sido suficientemente explorado en comparación con las investigaciones desarrolladas en español como lengua materna.

Si bien se trata de un corpus pequeño y, por lo tanto, los resultados deben interpretarse tomando muy en cuenta esta limitación, las muestras de habla obtenidas permiten establecer distintas variables relacionadas con el uso del léxico, detectar relaciones entre ellas y establecer conclusiones sobre la diversidad léxica. A continuación, se exponen estas conclusiones.

El discurso oral desarrollado por los estudiantes de ELE presenta un índice de diversidad inferior (0,47). Este resultado numérico adquiere relevancia si se relaciona con los descriptores de competencia léxica establecidos por el MCER para el nivel C1 (v. p. 16): riqueza de vocabulario y dominio de vocabulario.

A pesar de que se trata de hablantes de nivel C1 (se debe recordar que el nivel fue determinado por especialistas en el área), la riqueza de vocabulario correspondería más bien al descriptor del nivel B1:



B1: Tiene suficiente vocabulario para expresarse con algún circunloquio sobre la mayoría de los temas pertinentes para su vida diaria, como, por ejemplo, familia, aficiones e intereses, trabajo, viajes y hechos de la actualidad.

El descriptor correspondiente a B2 no resultaría el más apropiado en tanto enfatiza, nuevamente, la amplitud de vocabulario, aun cuando señala la presencia de vacilaciones:

B2: Dispone de un amplio vocabulario sobre asuntos relativos a su especialidad y sobre temas más generales. Varía la formulación para evitar la frecuente repetición, pero las deficiencias léxicas todavía pueden provocar vacilación.

Por otra parte, en cuanto al dominio de vocabulario, los resultados sí se corresponden con las expectativas del nivel C1, por cuanto no suelen presentarse “errores importantes de vocabulario”. Esta valoración se corrobora con los datos obtenidos en relación con el uso de palabras comodín o proformas léxicas, que, en términos globales, representa tan solo el 0,68 % del total de palabras del corpus. Asimismo, el uso de palabras en la lengua materna de los estudiantes se presenta de forma muy aislada (cuatro palabras en una de las cinco presentaciones), así como el caso de palabras utilizadas con un valor distinto al que tienen en la lengua meta (dos casos) y de peticiones de ayuda al profesor para cerciorarse de que la palabra seleccionada se ajusta al significado que se quiere expresar (seis casos).

Por lo tanto, de acuerdo con los resultados, la competencia léxica se ve afectada de forma negativa por la carencia de vocabulario variado. El bajo índice de diversidad léxica está relacionado con la falta de competencia en géneros discursivos propios del ámbito educativo, lo cual confirma los resultados obtenidos por Sánchez Avendaño (2005) y por Vargas Castro (2015) para la producción escrita en español como lengua materna y en ELE, respectivamente: 1) el discurso evidencia la falta de mecanismos para crear nuevas referencias sobre una misma unidad léxica y 2) esta falta parece obedecer más a la carencia de pautas para la construcción de un género discursivo en particular que



únicamente a insuficiencias relacionadas con el nivel de conocimiento y desempeño en la lengua meta.

Partiendo de que los géneros discursivos orales propios del ámbito educativo presentan características que pueden ser sistematizadas y descritas (Vilà, 2005) y de que, desde el punto de vista del proceso, el lenguaje es el producto de la integración de varios subsistemas –fonético, léxico, morfosintáctico, pragmático, discursivo– que poseen una autonomía relativa (Puyuelo, Rondal y Wiig, 2000, p. 18), se ha analizado de forma separada (no aislada, claro está) la diversidad léxica, con el fin de valorar su grado de desarrollo en presentaciones orales. Este análisis evidencia la necesidad de plantear y ejecutar actividades específicas que favorezcan el dominio de determinados tipos de discurso, según necesidades educativas en particular.

Resulta importante tener muy presente que la diversidad léxica, por un lado, y el conocimiento y el uso funcional de las palabras, por otro, deben valorarse de forma diferente, puesto que, en situaciones cotidianas, las palabras que utilizamos corresponden a un número bastante limitado, aproximadamente mil (Alvar, 2005). Esta premisa debe tomarse en cuenta puesto que implica la imposibilidad de valorar la diversidad léxica mediante una prueba de carácter general de dominio de la lengua. Tal como se evidenció, un porcentaje de diversidad léxica bajo no impidió que los estudiantes fueran capaces de cumplir con la tarea comunicativa de forma satisfactoria.

El porcentaje de diversidad léxica de 0,47 es considerado bajo tomando como referencia más importante el resultado obtenido por Barkelius (2007), dado que su estudio se centra en el discurso oral en ELE. En su caso, el porcentaje de diversidad léxica fue de 0,84 tanto en hablantes nativos como no nativos del español. Valga tomar en cuenta, eso sí, que su corpus consiste en narraciones basadas en las ilustraciones de un libro, y que, además, estas narraciones fueron segmentadas para incluir solamente las primeras 294 palabras, las cuales se dividieron entre sus correspondientes tipos.

Tanto para medir como para favorecer la diversidad léxica, componente central de la competencia léxica según los criterios tanto del MCER como del PCIC, se debería partir de una



caracterización del género discursivo que se pondrá en práctica, de la situación comunicativa en la que este se enmarca y del uso de estrategias que promuevan la variedad, la precisión y la propiedad léxicas acordes con los diferentes tipos de discurso.

Al respecto, el PCIC establece que este tipo de abordaje facilita que el alumno obtenga una visión tanto de la estructura global del discurso como de los distintos componentes que la conforman, partiendo de lo general a lo particular y viceversa, lo cual permite mantener una perspectiva contextualizada de la lengua.

Por lo tanto, existen varios aspectos que deben valorarse para establecer un marco general para abordar la diversidad léxica en el aula, los cuales se presentan en la Figura 1:

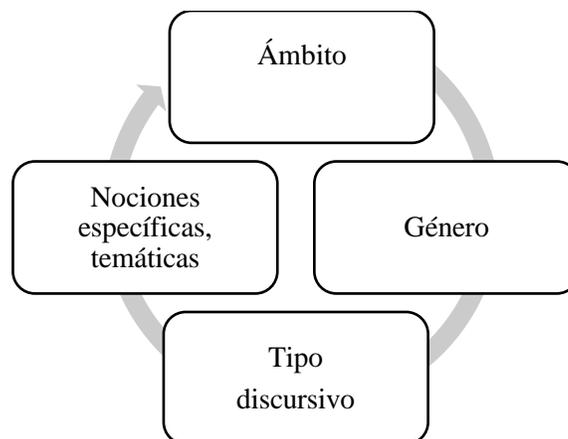


Figura 1. Aspectos que deben tomarse en cuenta para establecer un marco general para abordar la diversidad léxica en el aula
Fuente: Elaboración propia.

Siguiendo la propuesta del PCIC, el género constituye el punto de partida para la caracterización del discurso y corresponde a una forma de comunicación reconocida como tal por una comunidad de hablantes. En el caso del presente artículo, el género analizado correspondió a la presentación oral dentro del ámbito educativo, en particular, el universitario.



Los géneros se realizan mediante determinados tipos discursivos que consisten, a su vez, en manifestaciones textuales que presentan una organización prototípica según sus fines comunicativos. En este trabajo, la diversidad léxica se analizó a partir de un texto expositivo.

Por último, las nociones específicas se desarrollaron en las presentaciones orales a través de cinco temas: el cambio climático, la pobreza, la relación entre el poder económico y el Estado, la noción de tiempo como diferencia cultural y el impacto de la televisión en la sociedad actual.

Desde un punto de vista de ejecución en el aula, dado que los distintos aspectos incluidos en la Figura 1 interactúan entre sí, resulta indispensable establecer criterios de valoración de la diversidad léxica y brindarles a los estudiantes estrategias para utilizar un vocabulario variado acorde con las temáticas tratadas. De lo contrario, tal y como se ha evidenciado en este trabajo, los estudiantes de ELE recurrirán a su conocimiento funcional del léxico (tal como lo haría un hablante nativo) y serán capaces de cumplir con la tarea comunicativa planteada, pero sin que esto implique necesariamente un alto porcentaje de diversidad léxica.

Referencias bibliográficas

- Adam, J. M. (1992). *Les textes: types et prototypes. Récit, description, argumentation, explication et dialogue*. París, Francia: Armand Colin.
- Alvar, M. (2005). La frecuencia léxica y su utilidad en la enseñanza del español como lengua extranjera. www.cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/15/15_0017.pdf
- Araya, J. (2012). La competencia en la expresión oral de niños escolares en Costa Rica. El componente léxico. *Kañina, Revista de Artes y Letras*, 36 (1), 169-183.
- Ávila, R. (1986). "Léxico infantil de México: Palabras, tipos, vocablos". En *Actas del II Congreso Internacional sobre el español de América*. México, D. F.: Universidad Nacional Autónoma de México, 510-517.



- Ávila, R. (1988). Lengua hablada y estrato social. Un acercamiento lexicoestadístico. *NRFH*, 36 (1), 131-148.
- Barahona Novoa, A. (1996). *Léxico básico de la radio costarricense* [Tesis de Maestría en Lingüística]. Universidad de Costa Rica.
- Barkelius, J. (2007). Riqueza léxica y fluidez [Tesis de licenciatura]. Universidad de Estocolmo.
- Bassols, M. y Torrent, A. (2012). *Modelos textuales. Teoría y práctica* (3.ª ed.). Barcelona, España: Octaedro.
- Bellés Fortuño, B. y Fortanet Gómez, I. (2005). Spoken Academic Discourse: An approach to research on lectures. *Revista española de lingüística aplicada*, 1, 161-178.
- Berton, M. (2014). La riqueza léxica en la producción escrita de estudiantes suecos de ELE [Tesis de maestría]. Universidad de Estocolmo.
- Briz, A. (2000). *¿Cómo se comenta un texto coloquial?* Barcelona, España: Ariel.
- Calsamiglia, H. y Tusón, A. (2012). *Las cosas del decir. Manual de análisis del discurso*. Barcelona, España: Ariel.
- Camacho Adarve, M.^a M. (2009). *Análisis del discurso y repetición: palabras, actitudes y sentimientos*. Madrid, España: Arco Libros.
- Carroll, J. B. (1964). *Language and thought*. Nueva Jersey, EE. UU.: Prentice-Hall.
- Castellà, J. M. y Vilà, M. (2005). La lengua oral formal: características lingüísticas y discursivas. En M. Vilà (Coord.), *El discurso oral formal. contenidos de aprendizaje y secuencias didácticas* (pp. 25-36). Barcelona, España: Graó.
- Centro Virtual Cervantes. (2020). *Diccionario de términos claves de ELE*. https://cvc.cervantes.es/ENSEÑANZA/biblioteca_ele/diccio_ele/indice.htm
- Durán, P., Malvern, D., Richards, B. y Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics*, 25, 220-242.



- Gil del Moral, A. M. (2017). *Aprendizaje de español académico por alumnado de movilidad internacional en la Universidad de Alicante* [Tesis doctoral]. Universidad de Alicante.
- Haché, A. M. (1991). Aportes de las pruebas de riqueza léxica a la enseñanza de la lengua materna. En H. López Morales (Ed.), *La enseñanza del español como lengua materna* (pp. 47-60). Río Piedras, Puerto Rico: Universidad de Puerto Rico.
- Higueras García, M. (1997). Las unidades léxicas y la enseñanza del léxico a extranjeros. REALE. *Revista de Estudios de Adquisición de la lengua española*, 8, 35-50.
- Instituto Cervantes. (2002). *Marco Común Europeo de Referencia para las Lenguas: Aprendizaje, enseñanza, evaluación (MCER)*. Madrid, España: Grupo ANAYA, S. A.
- Instituto Cervantes. (2006). *Plan curricular del Instituto Cervantes. Niveles de referencia para el español. C1-C2*. Madrid, España: Biblioteca Nueva.
- Izquierdo, M. C. (2005). *La selección del léxico en la enseñanza del español como lengua extranjera*. Málaga, España: ASELE, Colección Monografías, N.º 8.
- Jarvis, S. (2013). Capturing diversity in lexical diversity. *Language Learning*, 63, 87-106.
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing.
https://www.researchgate.net/publication/285675449_Lexical_diversity_and_lexicaldensity_in_speech_and_writing_A_developmental_perspective.
- Justicia, F. (1995). *El desarrollo del vocabulario. Diccionario de frecuencias*. Granada, España: Universidad de Granada.
- Lara, L. F. (2006). *Curso de lexicología*. México: El Colegio de México.
- Laufer, B. (1991) The development of L2 lexis in the expression of advanced language learner. *Modern Language Journal*, 75, 440-448.
- Laufer, B. y Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16 (3), 307-322.



- Leoni de León, J. A. (1997). *Léxico básico de la prensa costarricense* [Tesis de licenciatura]. Universidad de Costa Rica.
- López Chávez, J. (1994). Comportamiento sintáctico de algunos verbos ordenados según su grado de disponibilidad léxica. *REALE, 1*, 67-84.
- López Morales, H. (1995-1996). Los estudios de disponibilidad léxica, pasado y presente. <https://boletinfilologia.uchile.cl/index.php/BDF/article/view/19231/20354>.
- López Morales, H. (2010). Los índices de riqueza léxica y la enseñanza de lenguas. https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/21/21_0015.pdf.
- Madrigal, M. y Vargas, E. (2016). Índice de riqueza léxica en redacciones escritas por estudiantes universitarios. *Káñina, Revista de Artes y Letras, 40* (Extraordinario), 139-147.
- Morales, A. (1986). *Léxico básico del español de Puerto Rico*. San Juan, Puerto Rico: Academia puertorriqueña de la lengua española.
- Müller, C. (1973). *Estadística lingüística*. Madrid, España: Gredos.
- Murillo Rojas, M. (1998). Riqueza léxica en niños preescolares costarricenses. *Revista Educación, 22* (1), 131-140.
- Murillo Rojas, M. (2009). Diversidad de vocabulario en los preescolares. Aportes para valorar su competencia léxica. *Revista de Filología, Lingüística y Literatura, 35* (1), 123-138.
- Puyuelo, M., Rondal, J. A. y Wiig, E. (2000). *Evaluación del lenguaje*. Barcelona, España: Masson.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge, Inglaterra: Cambridge University Press.
- Sánchez Avendaño, C. (2005). Los problemas de redacción de los estudiantes costarricenses: una propuesta de revisión desde la lingüística del texto. *Revista de Filología y Lingüística, 31* (1), 267-295.
- Terrádez, M. (2000). Frecuencias léxicas y análisis estadístico. En Briz, A. *¿Cómo se comenta un texto coloquial?* (pp.111-123). Barcelona, España: Ariel.



- Vargas Castro, E. (2015). Uso del léxico en textos escritos por estudiantes de nivel B2 de español como segunda lengua. *Revista de Filología, Lingüística y Literatura*, 41 (Número extraordinario), 81-92.
- Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17, 65-83.
- Vigara, A. M. (2005). *Morfosintaxis del español coloquial. Esbozo estilístico*. Madrid, España: Gredos.
- Vilà, M. (Coord.). (2005). *El discurso oral formal. Contenidos de aprendizaje y secuencias didácticas*. Barcelona, España: Graó.
- Villar, C. M. (2011). Las presentaciones académicas orales de los estudiantes alemanes de ELE: Del discurso monológico al dialógico. www.nebrija.com/revista-linguistica/las-presentaciones-academicas-oralesde-los-estudiantes-alemanes-de-e-le.
- Yu, G. (2010). Lexical diversity in MELAB writing and speaking task performances. http://wbaal.myzen.co.uk/proc07/32_guoxing_yu.pdf.



Esta obra está bajo una [licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)