

Complejidad, precisión, fluidez y léxico: Una revisión

IRINI MAVROU

Universidad Antonio de Nebrija

Resumen

El objetivo de este artículo es presentar y valorar el alcance de algunas aportaciones relevantes en el ámbito de Adquisición de Segundas Lenguas sobre los constructos de complejidad sintáctica y léxica, precisión lingüística y fluidez. Después de analizar brevemente su conceptualización en la bibliografía especializada, se presentan algunas de las medidas más frecuentemente empleadas en la operacionalización de estos constructos y se discuten las fortalezas y debilidades de la aplicación de dichas medidas en la evaluación de la actuación y del desarrollo interlingüístico de los aprendientes de lengua extranjera, con énfasis en el discurso escrito.

Palabras clave: Complejidad, precisión, fluidez, léxico

1. Introducción

Uno de los retos que han tenido que afrontar los investigadores en el ámbito de Adquisición de Segundas Lenguas (ASL) ha sido evaluar de manera objetiva los constructos multicomponentiales de la actuación y del desarrollo interlingüístico (Wolfe-Quintero, Inagaki, & Kim, 1998; Ellis & Larsen-Freeman, 2006; Larsen-Freeman, 2006, 2009; Housen & Kuiken, 2009; Housen, Kuiken, & Vedder, 2012). La elaboración de un índice de desarrollo [*developmental index*] que se tradujera en medidas específicas de la producción lingüística constituía una instancia ineludible y de gran valor práctico para diferenciar entre aprendientes con distinto grado de dominio de la lengua, posibilitar la obtención de datos comparables entre diferentes estudios y examinar la influencia en el uso de la lengua de factores, no necesariamente lingüísticos, que se implican en el proceso de aprendizaje y adquisición de una lengua extranjera (LE) (Wolfe-Quintero et al., 1998; Ellis & Larsen-Freeman, 2006). Gracias a las aportaciones de varios autores (véase Larsen-Freeman, 2009), las nociones de complejidad, precisión y fluidez figuraron como los principales componentes de la actuación y de la competencia lingüística del aprendiente de LE y emergieron en muchos estudios como “el principal epifenómeno de los mecanismos y procesos psicolingüísticos que subyacen a la adquisición, la representación y el procesamiento del conocimiento sobre la LE” (Housen & Kuiken, 2009:462).

A pesar de que estos constructos han sido ampliamente examinados en ASL, resulta oportuno mencionar que la investigación en este campo no ha estado exenta de desafíos, algunos de los cuales abarcan los siguientes aspectos: (1) su conceptualización como constructos científicos y su relación con los modelos teóricos de procesamiento y competencia lingüística en LE; (2) su medición empírica; (3) su in(ter)dependencia tanto en la actuación en un momento dado

como en el curso de desarrollo de la interlengua (IL); (4) la naturaleza de la relación de estas dimensiones con los mecanismos que subyacen a la adquisición, sean estos lingüísticos, psicolingüísticos o cognitivos; (5) los factores contextuales (características de las tareas, condiciones de implementación de estas), de personalidad (extraversión, ansiedad), cognitivos (aptitud, memoria operativa) y sociocognitivos (motivación) que afectan su manifestación sincrónica y diacrónica (Housen & Kuiken, 2009; Housen et al., 2012).

No es objeto del presente artículo dilucidar todas estas cuestiones, sino más bien recoger algunas aportaciones relevantes sobre la conceptualización de estos constructos en el ámbito de ASL y comentar algunos de los problemas metodológicos que plantea su medición empírica. Asimismo, dada la amplitud del tema, el presente artículo se centrará principalmente en la evaluación de la expresión escrita, una destreza de suma importancia en el ámbito académico y profesional, aunque en ocasiones desestimada dentro de los enfoques comunicativos de la enseñanza de LE.

2. Complejidad lingüística

La complejidad, caracterizada como un “palimpsesto” por Housen et al. (2012), constituye el constructo más problemático y ambiguo (Housen & Kuiken, 2009; Palloti, 2009; Bulté & Housen, 2012; Housen et al., 2012). La naturaleza multidimensional y la polisemia de esta noción (Palloti, 2009; Bulté & Housen, 2012), las diferencias estilísticas entre los individuos (Palloti, 2009) – especialmente evidentes en el caso de la expresión escrita– y la influencia del tipo de tarea en los niveles de complejidad exhibidos (véase, por ejemplo, Skehan & Foster, 1997; Foster & Tavakoli, 2009) ponen en evidencia la dificultad que supone llegar a una definición unánime del constructo.

La complejidad lingüística se ha concebido, por lo general, en términos de elaboración, variación y sofisticación de las estructuras lingüísticas empleadas por el aprendiente de LE (Wolfe-Quintero et al., 1998). No obstante, tales términos se han caracterizado como “muy generales, vagos o incluso circulares” (Bulté & Housen, 2012:22). Como observan muy acertadamente Bulté y Housen (2012), en la mayoría de los estudios, la complejidad lingüística se suele conceptualizar de una manera intuitiva y en el nivel más bajo de especificación del constructo [*operational-statistical construct*], esto es, únicamente en función de las medidas cuantitativas utilizadas para su operacionalización. Frente a esta tendencia reduccionista, Bulté y Housen (2012) subrayan la necesidad de una caracterización más explícita de esta noción.

Según la propuesta de estos autores, la complejidad consta de al menos tres componentes: (1) la complejidad proposicional [*propositional complexity*] que concierne al número de ideas y unidades de información, (2) la complejidad discursiva-interaccional [*discourse-interactional complexity*] que se refiere al número y tipo de cambios de turno [*turn changes*], a los movimientos interaccionales y a los papeles de participación adoptados durante el discurso

dialógico, y (3) la complejidad lingüística [*linguistic complexity*] (Bulté & Housen, 2012), que es el tipo de complejidad que interesa en este artículo.

En cuanto a la complejidad lingüística, se ha interpretado de dos maneras diferentes: (1) como una propiedad dinámica del sistema de IL, que se refiere al grado de elaboración, la riqueza y la diversidad de este sistema [*global* o *system complexity*] y (2) como una propiedad más estable de elementos, estructuras o reglas lingüísticas que conforman el sistema de IL [*local* o *structure complexity*] entendida más bien en términos de profundidad (Housen & Kuiken, 2009; Bulté & Housen, 2012).

Palloti (2009), sin embargo, plantea la siguiente pregunta: ¿hasta qué punto es apropiado tratar los aspectos de la complejidad identificados en la bibliografía especializada como diferentes manifestaciones del mismo constructo o como si fueran constructos diferentes? Bulté y Housen (2012), por ejemplo, asumen que los tipos de complejidad propuestos por ellos son constructos diferentes solo en la teoría y, probablemente, mucho más enlazados en contextos reales de aprendizaje.

La conceptualización de la complejidad sintáctica y de la léxica como aspectos pertenecientes a la misma o a diferentes áreas de la actuación constituye también para Skehan (2009) una cuestión abierta al debate. Según el autor, la selección de los elementos léxicos que formarán parte del discurso no causa demasiados problemas a los hablantes nativos, de ahí que la complejidad adquiera “un carácter más unidimensional” en el caso de la lengua materna (LM); no obstante, para los aprendientes de una LE las dos dimensiones lingüísticas, complejidad léxica y sintáctica, “no parecen integrarse tan bien” (Skehan, 2009:528). Es por ello que Skehan (2009) subraya la necesidad de incluir en los estudios de ASL medidas de léxico –un área relegada en la investigación empírica de tareas–, ya que “el léxico representa una forma de complejidad” (Skehan, 2009:514).

2.1 Complejidad sintáctica

La complejidad sintáctica se define como la habilidad del aprendiente de LE de utilizar una variedad de formas y estructuras sintácticas, tanto básicas como más sofisticadas, a la hora de desempeñar tareas que requieren la producción del lenguaje (Wolfe-Quintero et al., 1998; Ortega, 2003). Asimismo, esta dimensión lingüística se suele vincular con la voluntad de experimentar y comprobar hipótesis acerca del sistema de la lengua y de asumir riesgos en busca de enunciados nuevos (Skehan & Foster, 2001).

Hunt (1965) propuso unos índices que permitiesen la cuantificación del desarrollo sintáctico de la expresión escrita de niños, siendo los más utilizados los índices primarios que se basan en dos unidades de análisis, la unidad terminal o unidad t (UT) y la cláusula (CL). Checa García (2005), basándose en los presupuestos teóricos de Hunt, hizo las siguientes puntualizaciones respecto a las etiquetas utilizadas (Checa García, 2005):

- Cláusula (CL): una oración simple o una proposición, principal o subordinada, con o sin sujeto expreso; se incluyen formas no personales del verbo, salvo que funcionen en perífrasis verbales
 - Longitud de la unidad terminal (LUT): palabras (PAL) / unidades t (UT)
 - Longitud de la cláusula (LCL): palabras (PAL) / cláusulas (CL)
 - Índice de subordinación (IS): cláusulas (CL) / unidades t (UT)
 - $LUT=IS \times LCL$ (Hunt, 1965; Checa García, 2005)
 -

A pesar de que muchos autores han sido partidarios de los presupuestos de Hunt y han incluido en sus estudios medidas como la LUT, la LCL y el IS, debe advertirse que dichas medidas no han estado exentas de críticas. De hecho, se establecieron nuevas medidas, mientras que otras se refinaron o se modificaron, dependiendo de los objetivos del investigador y del grupo meta del estudio.

En primer lugar, es indispensable ofrecer algunas aclaraciones sobre las unidades de análisis de la complejidad sintáctica. Según Skehan (2003), la UT es la más apropiada para el discurso escrito. Ishikawa (1995), sin embargo, argumentó a favor de la CL, que es una unidad más pequeña, para el análisis del discurso escrito de aprendientes principiantes. Otras unidades de análisis de la complejidad sintáctica que han sido propuestas son la unidad C (C-unit) (véase Crookes, 1990) y la unidad *Analysis of Speech* (AS-unit) (Foster, Tonkyn, & Wigglesworth, 2000).

La segunda cuestión que resulta oportuno considerar es la controversia sobre la definición del concepto de CL (Wolfe-Quintero et al., 1998; Bulté & Housen, 2012). En español, por ejemplo, se pueden identificar cuatro tipos de CL: independiente, adjetiva/relativa, adverbial y sustantiva/nominal; además, los tres últimos tipos pueden reducirse a proposiciones de participio, gerundio e infinitivo. Siguiendo el ejemplo de Wolfe-Quintero et al. (1998), mientras que para Hunt (1965) la CL comprende cuatro tipos, para aquellos autores que optan por definiciones de la CL similares a la de Checa García (2005) el mismo concepto se refiere a siete tipos –incluidos los participios, los gerundios y los infinitivos–, y esto dará valores diferentes tanto del IS como de la LCL.

Sea cual sea el modo de operacionalizar el concepto de CL, en ambos casos habrá ventajas e inconvenientes. Como explican Bulté y Housen (2012), la primera es una definición operativa pero con un alcance limitado, mientras que la segunda, aunque lingüísticamente más válida, puede arrojar valores de la LCL poco representativos. De lo anterior se hace patente, por un lado, la responsabilidad que recae en el investigador de explicitar con el mayor grado de detalle posible lo que concibe como CL y, por otro lado, la necesidad de considerar diferentes conceptualizaciones a la hora de contrastar los resultados de su estudio.

Respecto a las medidas utilizadas en la evaluación del constructo, se pueden identificar dos categorías amplias, las medidas de longitud promedio de las unidades UT, CL, C-unit, AS-unit y oración (OR) y las medidas que evalúan la

densidad de subordinación (IS, CL/C-unit, CL/AS-unit, CL/OR, CLdep/UT, CLdep/CL; CLdep=cláusulas dependientes).

La primera categoría ha suscitado opiniones divergentes. A pesar de que en algunos estudios (Sotillo, 2000; Kormos, 2011; Vajjala & Meurers, 2012; Treffers-Daller, 2013) las medidas de longitud fueron utilizadas como indicadores de la complejidad sintáctica, otros autores (Wolfe-Quintero et al., 1998; Torras & Celaya, 2001; Navés, Torras, & Celaya, 2003; Larsen-Freeman, 2006; Nariman-Jahan & Rahimpour, 2011; Yang & Sun, 2015) las trataron como propias de la fluidez. Un punto de vista diferente mantiene Navés (2007) al sostener que las medidas LCL y longitud promedio de la oración deberían considerarse representativas de un constructo diferente y todavía por determinar. Bulté y Housen (2012), por su parte, caracterizaron como ambiguas o híbridas algunas medidas, en el sentido de que evalúan simultáneamente varias facetas de la complejidad. En concreto, según estos autores, medidas como la LUT podrían considerarse como genéricas de la complejidad sintáctica o incluso de la competencia lingüística global (Bulté & Housen, 2012).

Teniendo como telón de fondo tales disyuntivas y lejos de reiterar una posición a favor o en contra del uso de las medidas de longitud como indicadores de la complejidad sintáctica, de la fluidez o de otro constructo, es necesario subrayar la necesidad de investigaciones futuras que se centren en probar tales supuestos mediante la aplicación de técnicas estadísticas multivariantes como el Análisis Factorial Exploratorio (AFE).

Respecto a las medidas de subordinación, medidas por defecto de la complejidad sintáctica, la problemática va por otro camino. Como explican Bulté y Housen (2012), a estas medidas se les suele dar mayor peso a la hora de examinar la complejidad sintáctica, ya que parecen plantear un mayor reto en términos de procesamiento cognitivo. No obstante, como advierten varios autores (Palloti, 2009; Bulté & Housen, 2012), si bien una oración gramaticalmente compleja puede requerir más recursos cognitivos, la correspondencia entre complejidad cognitiva y complejidad sintáctica es un asunto que debería demostrarse empíricamente.

Llegados a este punto, es necesario precisar lo siguiente: (a) mayor complejidad sintáctica o subordinación no significa necesariamente mayor desarrollo o mayor calidad del discurso (Ortega, 2003); (b) un mayor grado de elaboración sintáctica puede ser el resultado de variables como el conocimiento sobre los mecanismos sintácticos de la lengua que se aprende (Checa García, 2005), la transferencia del estilo propio de escritura en la LM u otra lengua dominante a la LE o incluso la forma de enseñar este aspecto lingüístico.

En síntesis, la complejidad sintáctica constituye un constructo multifacético que debería estudiarse teniendo en cuenta los diferentes niveles de su manifestación y mediante el empleo de medidas complementarias que permitan evitar la redundancia en la medición (Bulté & Housen, 2012). Además, estas medidas deberían seleccionarse de acuerdo con el nivel de competencia lingüística del grupo meta, las características de las tareas implementadas y tras proporcionar una

definición lo más clara y explícita de las unidades de análisis eventualmente utilizadas.

2.2 Complejidad léxica

El léxico ha sido la dimensión que menos presencia ha tenido (Skehan, 2009) en lo que se ha denominado el estudio de CAF [*Complexity, Accuracy, Fluency*]. Ante todo, es pertinente aclarar que aquí interesa el vocabulario productivo del aprendiente de LE y se utiliza el término riqueza (Laufer & Nation, 1995; Lu, 2012; Šišková, 2012) o complejidad léxica (Wolfe-Quintero et al., 1998; Skehan, 2009; Bulté & Housen, 2012) que se concibe como un constructo multidimensional compuesto por al menos tres aspectos: la diversidad o variación léxica, la sofisticación léxica y la densidad léxica (Laufer & Nation, 1995; Wolfe-Quintero et al., 1998; Bulté & Housen, 2012; Lu, 2012).

En cuanto al primer aspecto, las innumerables definiciones propuestas en la bibliografía especializada, utilizadas de manera “intercambiable, jerárquica o excluyente” (Yu, 2007:81), han convertido la diversidad léxica en un constructo complejo no solo en su conceptualización, sino también en su medición (McCarthy & Jarvis, 2007, 2010; Yu, 2007). Siguiendo a McCarthy y Jarvis (2007, 2010), la diversidad léxica se entiende como la cantidad y variedad del vocabulario utilizado, tal y como se refleja en el uso de la lengua. Partiendo de esta definición, es posible hacer una distinción más precisa, considerando, por una parte, la cantidad de ítems diferentes [*word types*], esto es, la diversidad, y, por otra parte, la variación de estos ítems en relación con ciertas propiedades del vocabulario (por ejemplo, verbos y palabras léxicas, es decir, de contenido referencial y semántico). En ambos casos, un alto grado de diversidad o variación lleva implícita una menor repetición (Johansson, 2008).

La identificación de un índice fiable de la diversidad léxica ha sido especialmente problemática (McCarthy & Jarvis, 2010). La medida convencional de *type/token* ratio (TTR) ha sido criticada por varios autores (Richards, 1987; Wolfe-Quintero et al., 1998; Richards & Malvern, 2000; McCarthy & Jarvis, 2007, 2010; Skehan, 2009; Treffers-Daller, 2013) debido a que se presenta sensible a la longitud de los textos comparados (textos largos tienden a dar valores pequeños de TTR).

A fin de compensar este problema se han propuesto diferentes procedimientos de normalización, aunque tampoco han aportado la solución esperada. La extracción de un número fijo de palabras (consecutivas o seleccionadas aleatoriamente) o el truncamiento de los textos a una longitud establecida, si bien ofrecen un valor aproximado de la diversidad léxica, hacen que se pierda mucha información de los datos (Laufer & Nation, 1995; Richards & Malvern, 2000; Lu, 2012). Teniendo en cuenta, además, que diferentes partes del discurso suelen presentar distintos grados de diversidad léxica, optar por procedimientos de selección aleatoria puede acarrear el riesgo de incluir en el análisis una muestra no representativa del discurso de un hablante o escritor (McCarthy & Jarvis, 2010). Wolfe-Quintero et al. (1998), por su parte, recomiendan las transformaciones

algebraicas de la TTR, siempre que se imponga un límite de tiempo para el desempeño de las tareas o que se trate de tareas del mismo tipo y de características similares (límite conceptual).

Otra medida frecuentemente utilizada en la investigación empírica de ASL ha sido la *Mean Segmental Type Token Ratio*. Su valor se obtiene tras dividir el discurso en segmentos de una longitud dada (50 o 100 palabras) y calculando, a continuación, la TTR promedio de estos segmentos. No obstante, este procedimiento acarrea los siguientes problemas: (a) a veces resulta imposible dividir las muestras del discurso en segmentos de igual tamaño; (b) diferentes autores optan por diferentes criterios de segmentación; (c) dicha medida no toma en cuenta la repetición de palabras entre diferentes segmentos (Richards & Malvern, 2000).

En cuanto a las transformaciones algebraicas de la TTR, si bien constituyen una manera alternativa para compensar la variabilidad entre textos de diferente longitud, tampoco están libres de problemas metodológicos. La TTR corregida ($\text{Types}/\sqrt{2\text{PAL}}$) y el índice de Guiraud ($\text{Types}/\sqrt{\text{PAL}}$), aunque toman en cuenta la presencia de muchas formas en una composición larga, no son medidas ‘puras’, dada la transformación aplicada al número de *tokens*, por lo que, como se señaló anteriormente, su uso sería deseable siempre que haya un límite temporal o conceptual (Wolfe-Quintero et al., 1998).

En lo que respecta a la fórmula D (calculada a través del programa CLAN en CHILDES; MacWhinney, 2000), medida que se considera poco sensible a la longitud de los textos, trabajos recientes vienen a cuestionar su efectividad. Yu (2007), por ejemplo, aunque reconoce la superioridad de la D en comparación con otras medidas más convencionales, sostiene que esta resulta más adecuada para evaluar la diversidad léxica del discurso oral. Treffers-Daller (2013) apunta que la utilidad de la D puede verse limitada cuando se aplica a la comparación de idiomas diferentes, ya que las posibles fluctuaciones en los valores de esta medida pueden deberse a las diferencias tipológicas entre las lenguas estudiadas, conclusión a la que también llegaron Woerfel y Yilmaz (2011). Sin embargo, estas discrepancias pueden evitarse si los datos se lematizan apropiadamente (Treffers-Daller, 2013) y, según parece, esto tiene vigencia no solo para el cómputo de la D, sino también para otros índices basados en la relación *type/token*.

Otro dato interesante proviene de los estudios de McCarthy y Jarvis (2007, 2010). En McCarthy y Jarvis (2007) se compararon trece medidas de diversidad léxica y se comprobó que la decisión sobre la más apropiada debe basarse en el rango de palabras que comprenden las muestras del discurso comparadas. En un estudio posterior, McCarthy y Jarvis (2010) demostraron la validez y la menor sensibilidad al problema de la longitud del discurso de dos relativamente nuevas aproximaciones a la diversidad léxica, la *Measure of Textual Lexical Diversity* (MTLD) y el índice HD-D. De hecho, Koizumi (2012), tras comparar la TTR, el índice de Guiraud, la D y la MTLD, encontró que la D no es adecuada para analizar textos breves, mientras que la MLTD requiere un mínimo de 100 *tokens* para que su cómputo sea fiable.

En función de lo expuesto, queda claro que una decisión arbitraria a la hora de seleccionar las medidas de diversidad léxica podría invalidar los resultados de un estudio. Asimismo, hay que tener en cuenta que la diversidad léxica es un constructo que no se manifiesta de manera aislada, sino que puede variar dependiendo de la modalidad del discurso, el tópico o el tipo de instrucciones dadas, aun habiéndose controlado factores como las habilidades de escritura o el nivel de competencia lingüística en la LE (Yu, 2007). Ante la pregunta de McCarthy y Jarvis (2007) sobre si una sola medida podría abarcar el constructo de diversidad léxica, resulta interesante la sugerencia de Lu (2012) de elegir un conjunto de medidas que tengan un impacto significativo en la calidad del discurso pero que no correlacionen fuertemente entre sí.

Las medidas comentadas anteriormente distinguen entre *tokens* (palabras) diferentes, pero no captan ni su tipo ni tampoco su distribución. Otras medidas para tal objetivo son las de variación léxica o verbal. Estas se aplican a diferentes clases de palabras y sus transformaciones algebraicas se consideran más apropiadas para examinar el constructo en cuestión (Wolfe-Quintero et al., 1998; Yu, 2007). No obstante, tampoco estas medidas son más informativas que sus equivalentes en lo que concierne a la calidad del discurso o la dificultad de las palabras empleadas, lo que condujo a varios autores, entre ellos, Laufer y Nation (1995) y Meara y Bell (2001), a optar por medidas un tanto diferentes, aunque también basadas en ratios: las medidas de sofisticación léxica.

El constructo de sofisticación o rareza léxica informa sobre la proporción de palabras poco comunes utilizadas en un texto (Laufer & Nation, 1995; Lu, 2012; Šišková, 2012) y la evaluación del vocabulario se hace en este caso con referencia a criterios externos (listas de frecuencias de palabras, listas de vocabulario básico o sofisticado). La medida convencional de sofisticación léxica ha sido la ratio de formas sofisticadas en relación con el número total de formas. A partir de allí, se han propuesto varias transformaciones de dicha ratio que, además de dar una indicación de la dificultad de los ítems empleados, permiten controlar el efecto de la longitud de los textos, de ahí que se consideren procedimientos metodológicamente más adecuados (Wolfe-Quintero et al., 1998; Lu, 2012; Šišková, 2012).

La principal dificultad a la hora de optar por una medida de sofisticación léxica reside en la definición del vocabulario básico o avanzado y, por consiguiente, en el criterio externo que servirá de punto de referencia. Según Laufer y Nation (1995), tal definición depende del punto de vista del investigador, quien deberá tomar en cuenta el nivel de dominio lingüístico del grupo meta. También, es preciso señalar que las medidas de sofisticación léxica son ratios, por lo que resulta imprescindible la lematización del corpus, al menos cuando se examinan lenguas con mucha flexión (Treffers-Daller, 2013).

Por último, la densidad léxica se corresponde con la ratio de ítems léxicos en relación con el número total de *tokens* o de *tokens* funcionales (preposiciones, interjecciones, pronombres, etc.) (Laufer & Nation, 1995; Wolfe-Quintero et al., 1998; Johansson, 2008; Lu, 2012; Šišková, 2012). Esta ratio ofrece una noción del

information packaging (Johansson, 2008). Dicho de otro modo, un texto con alta proporción de ítems léxicos debería contener más información que uno con alta proporción de ítems funcionales (Johansson, 2008).

Sin embargo, tales afirmaciones tienen que tomarse con precaución, ya que la ratio de la densidad léxica puede verse afectada por la edad de los aprendientes (Johansson, 2008), así como por las propiedades sintácticas y cohesivas de la lengua que se examina (Laufer & Nation, 1995; Wolfe-Quintero et al., 1998). Por ello, Laufer y Nation (1995) sostienen que esta ratio no mide necesariamente el léxico, sino que parece reflejar las características estructurales de una composición.

La cuestión se complica cuando se comparan grupos de aprendientes con diferente grado de dominio de la LE. Siguiendo las reflexiones de Wolfe-Quintero et al. (1998), en un nivel principiante los aprendientes suelen utilizar proporcionalmente un mayor número de ítems léxicos a fin de transmitir su mensaje. Si bien este grupo cuenta con un vocabulario bastante limitado, la ratio de la densidad léxica que obtendrá podría ser superior a la de un grupo de aprendientes competentes o de hablantes nativos cuyo discurso, como es de esperar, contiene más repeticiones y un mayor número de elementos gramaticales. Ello hizo concluir a Wolfe-Quintero et al. (1998) que la densidad léxica, a diferencia de la diversidad y la sofisticación, no se vincula con el desarrollo de la IL.

Para resumir, mediante las líneas anteriores se ha perseguido rescatar la importancia del constructo de complejidad léxica en la bibliografía sobre CAF. En cuanto a las medidas presentadas, es importante señalar que evalúan solo la amplitud del vocabulario y, por tanto, ignoran factores importantes del léxico como la función gramatical o la precisión de una palabra dentro de un contexto particular (Šišková, 2012). Aun así, su utilidad está fuera de dudas. Como bien apuntan McCarthy y Jarvis (2010), lo más importante es que el investigador utilice diferentes medidas potencialmente relevantes e informativas para el constructo que examina, que conozca bien sus limitaciones y que se responsabilice de las decisiones que tome.

3. Precisión lingüística

La precisión lingüística se ha caracterizado por varios autores como el constructo más antiguo y transparente y el que mayor consistencia interna presenta (Wolfe-Quintero et al., 1998; Housen & Kuiken, 2009; Palloti, 2009; Housen et al., 2012). El término se suele vincular con la habilidad de producir lengua carente de errores y el grado de conformidad con la norma, que suele ser la de un hablante nativo ideal de esta lengua (Wolfe-Quintero et al., 1998; Skehan & Foster, 2001; Housen & Kuiken, 2009; Palloti, 2009; Housen et al., 2012). Skehan y Foster (2001; Skehan, 1996) sostienen que la precisión puede reflejar un grado de conservadurismo, en el sentido de que el aprendiente de LE, en lugar de experimentar con la lengua, intenta evitar aquellas estructuras y elementos lingüísticos que puedan inducir al error y ello conduce a *trade-offs* entre la

precisión y la complejidad lingüística (para una posición diferente véase la Hipótesis de Cognición; Robinson, 2001, 2003).

Respecto a la medición del constructo se pueden identificar dos aproximaciones. La primera se centra en la precisión de cierto tipo de unidades de análisis y las medidas utilizadas son ratios con numerador el número total de UT libres de errores (UTLLE) o de CL libres de errores (CLLE) y denominador el número total de UT, CL, OR o PAL. La segunda aproximación se basa en la contabilización de los errores y, más precisamente, en la proporción de errores totales o específicos (morfológicos, sintácticos o léxicos) en relación con el número total de UT y CL o proporcionalmente a la cantidad total de palabras.

Como apuntan Wolfe-Quintero et al. (1998), las medidas de precisión no parecen estar relacionadas con el desarrollo interlingüístico, sino con el constructo de error. Palloti (2009) está de acuerdo con esta posición al señalar que es posible encontrar textos correctos pero poco desarrollados, así como textos con muchos errores que presenten varios rasgos de evolución; por tanto, precisión y desarrollo son constructos diferentes y, como tales, deberían ser evaluados mediante medidas diferentes.

En cuanto a las medidas que se basan en la UTLE, Hirano (1989) subraya la dificultad que supone la definición de la UTLE, y por consiguiente de la CLLLE, puesto que depende en gran medida del concepto de error. En otras palabras, por UTLE o CLLLE se puede concebir toda UT o CL con sentido, sin considerar problemas de ortografía o puntuación, o se puede considerar incorrecta toda UT o CL que contenga errores de sintaxis, léxico, ortografía y puntuación, es decir, cualquier tipo de error puede considerarse como fallo. Iwashita, Brown, McNamara y O'Hagan (2008), por su parte, señalan que la identificación de todo tipo de error, a pesar de ser un procedimiento más comprensivo, puede dificultar la obtención de un alto grado de consistencia interevaluador.

De lo anterior se desprende que, en los estudios que buscan cuantificar la corrección lingüística, es preciso y hasta cierto punto obligatorio proporcionar una definición detallada de este concepto. Además, la decisión de incluir o excluir cierto tipo de errores debería tomarse en función del nivel de dominio lingüístico del grupo meta y el valor discriminatorio de los errores dentro de la población estudiada (Wolfe-Quintero et al., 1998). Ahora bien, aun explicitando lo que se entiende por error, ninguno de los enfoques de medición de la precisión lingüística puede captar la gravedad de este, de ahí que las medidas presentadas anteriormente no sean suficientes por sí solas para evaluar el grado de corrección global de un texto escrito (Wolfe-Quintero et al., 1998).

A estas reflexiones hay que agregar los problemas que conlleva la evaluación de la expresión escrita de aprendientes de bajo nivel de competencia lingüística por medio de la UT, puesto que este grupo todavía no ha automatizado algunos recursos léxico-sintácticos necesarios para la producción de fragmentos comprensibles (Hirano, 1989). A modo de ejemplo, se reproducen a continuación dos oraciones:

- (1) En la primera imagen ver una mujer entrar la cocina.
- (2) En la primera viñeta veo a una mujer quiere entrar en la cocina.

La oración (1) plantearía dificultades a la hora de identificar la UT, aunque se entienda cuál es la intención comunicativa del hablante o escritor. Producciones en las que abunde el uso de infinitivos podrían ser típicas del discurso de un aprendiente de LE que se encuentra en un nivel principiante. En cuanto a la oración (2), que es más acertada y mejor desarrollada, nos encontramos frente al siguiente problema. Parece ser que la intención del hablante era producir la siguiente frase: ‘En la primera viñeta veo a una mujer que quiere entrar en la cocina’. En este caso tendríamos una UTLE. Sin embargo, la misma frase puede expresarse mediante dos oraciones: ‘En la primera viñeta veo a una mujer. [La mujer] Quiere entrar en la cocina’. Esto equivale a dos UTLE. Dilucidar estas cuestiones metodológicas requiere que los investigadores proporcionen toda la información relativa al procedimiento seguido, sea para la segmentación de las muestras del discurso en las UT y CL o para la cuantificación de los errores (véase también Polio, 1997).

Unsworth (2008) realiza algunas observaciones adicionales útiles de mencionar aquí. Como argumenta la autora, es posible que los aprendientes produzcan ciertos errores una y otra vez. Ante la presencia de errores repetidos se podría contabilizar solo la primera instancia. Sin embargo, al ignorar el mismo error en la segunda o tercera instancia, se le atribuiría al aprendiente conocimiento que no tiene. En caso contrario, es decir, si un error se contabiliza inicialmente como tal, el aprendiente no será recompensado si produce la misma forma correctamente. Por tanto, Unsworth (2008) concluye que los errores repetidos deberían considerarse por separado.

Una última consideración concierne al criterio según el cual se evalúa la precisión lingüística (Housen & Kuiken, 2009; Palloti, 2009; Housen et al., 2012). En líneas generales, este criterio suele basarse en lo que un hablante ideal de la lengua es capaz de producir, esto es, la norma prescriptiva estándar. Sin embargo, la norma no estándar (del hablante no nativo) puede ser igual de aceptable (véase Cook, 1999) y esto tiene vigencia para la evaluación no solo de la precisión, sino también de los demás constructos (complejidad sintáctica, complejidad léxica y fluidez). Asimismo, hay que tener en cuenta que lo que es gramaticalmente correcto en una variante puede no serlo en otra (por ejemplo, en el caso del inglés británico y el americano) y que un hablante nativo puede cometer errores incluso cuando se expresa en su LM (Polio, 1997). Es por ello que Housen et al. (2012) recomiendan que se interprete la precisión en términos de adecuación o aceptabilidad.

La adecuación es un aspecto en el que también se detiene Palloti (2009). El autor defiende la necesidad de aproximarse a la adecuación –y, por consiguiente, al grado de éxito en la consecución del objetivo comunicativo– como una dimensión distinta de la actuación, que, aunque independiente en teoría, a nivel

empírico podría presentar una mayor relación con las dimensiones lingüísticas de complejidad, precisión y fluidez.

En función de lo expuesto, se ha de hacer hincapié en la necesidad de disponer de una serie de criterios claros antes de proceder a la evaluación de la precisión lingüística del discurso de los aprendientes de una LE. La definición explícita del concepto de error y la especificación de la normativa que servirá de punto de referencia para el análisis constituyen un requisito imprescindible a fin de aumentar la replicabilidad de la metodología y facilitar el contraste de los resultados. Estas decisiones, sin embargo, no se deben tomar al margen de factores como el nivel de dominio lingüístico del grupo meta, las características de las tareas implementadas, ciertas propiedades de la lengua que se examina o incluso la importancia adjudicada por los propios alumnos a cierto tipo de errores. Por último, es preciso señalar que, si bien se ha constatado que algunas medidas genéricas como las UTLE/UT y Errores/UT resultan más apropiadas para evaluar la precisión lingüística (Wolfe-Quintero et al., 1998; Larsen-Freeman, 2006), no es solo recomendable, sino también necesario, complementarlas con otras más específicas (Robinson, Cadierno, & Shirai, 2009) y tras considerar diferentes opciones de análisis respecto a casos particulares de errores.

4. Fluidez escrita

La misma incongruencia terminológica de varios constructos en el campo de ASL atañe también al constructo de fluidez (Schmidt, 1992; Chambers, 1997; Housen & Kuiken, 2009; Housen et al., 2012). De hecho, en la bibliografía especializada se pueden encontrar conceptualizaciones que abarcan varios aspectos o conceptualizaciones más reducidas, las que incluyen dimensiones del constructo susceptibles de cuantificación y otras de carácter cualitativo, así como definiciones más lingüísticas o más cognitivas.

El término fluidez se ha utilizado de manera general como sinónimo de un alto dominio de la lengua y se ha vinculado con la velocidad, la elocuencia y la facilidad en la producción del lenguaje (Chambers, 1997). Para Skehan (1996; Skehan & Foster, 2001), la fluidez se corresponde con la atención al significado y se manifiesta en la capacidad de producir lengua a una velocidad normal, sin muchas interrupciones, apoyándose en el sistema basado en fragmentos lingüísticos lexicalizados almacenados en la memoria [*memory-based system*].

Schmidt (1992) habla de la fluidez como una habilidad procedimental automática [*automatic procedural skill*] que no requiere mucha atención o esfuerzo. Tras una revisión de diferentes teorías y modelos psicolingüísticos que hace el autor, se deja entrever que la fluidez podría desarrollarse mediante mecanismos cognitivos tales como los procesos de segmentación [*chunking*], la codificación de instancias y su recuperación desde la memoria [*memory retrieval mechanisms*] o el fortalecimiento progresivo de conexiones entre estímulo y respuesta que se consigue a través de la práctica, aunque ninguno de estos mecanismos por sí solo parece proveer una base sólida para explicar suficientemente el complejo proceso de desarrollo de la fluidez en una LE.

Desde una perspectiva diferente, Segalowitz (2007) distingue dos aspectos de la ‘fluidez cognitiva’. El primero se denomina *access fluidity* y es el proceso de vinculación de palabras y expresiones con sus correspondientes significados, mientras que el segundo se refiere al control de la atención [*attention control*], proceso mediante el cual el usuario de la lengua enfoca y reenfoca su atención al mensaje que va produciendo en tiempo real. Según Segalowitz (2007), el desarrollo de la fluidez se consigue a través de una amplia exposición y práctica con la lengua meta en situaciones comunicativas naturales.

Las líneas precedentes demuestran que existen varias maneras de abordar el concepto de fluidez y justifican su caracterización como un constructo multidimensional. Aspectos del constructo que han sido examinados en la investigación empírica de ASL comprenden variables temporales –velocidad, número, longitud y ubicación de las pausas–, ciertos fenómenos que contribuyen a la percepción de un lenguaje que se aproxima al de un hablante nativo como las autocorrecciones, los falsos comienzos, las repeticiones y otros fenómenos de vacilación o reparación, aspectos fonológicos como el acento y la entonación, así como variables más globales como la precisión (Kormos & Dénes, 2004; Skehan 2009; Götz, 2013).

No obstante, definiciones como las anteriores necesitan una reconceptualización a fin de poder cobrar sentido para el caso particular de la expresión escrita. El término fluidez escrita en sí mismo resulta ambiguo, ya que la producción escrita se entiende como un proceso lento, lleno de pausas y retrocesos (para revisar, reflexionar, reformular, corregir), un proceso solitario, durante el cual el escritor suele tener la plena libertad de buscar y apoyarse en fuentes de información externas.

Varios autores coinciden en afirmar que la bibliografía especializada carece de una definición viable de la fluidez escrita y, si la hay, la misma resulta bastante limitada (Bruton & Kirby, 1987; Fellner & Apple, 2006; Abdel Latif, 2008). Además, como apunta Abdel Latif (2008), las definiciones de la noción de fluidez escrita suelen reflejar diferentes maneras de operacionalizar el constructo, observación similar a la de Bulté y Housen (2012) sobre el constructo de complejidad lingüística.

La conceptualización de Bruton y Kirby (1987), sin embargo, ofrece algunos elementos interesantes. Adoptando una perspectiva evolutiva, estos autores diferencian entre una etapa inicial de la fluidez escrita que incorpora aspectos como la cantidad y la velocidad, un sentimiento de comodidad que permite al individuo escribir de forma rápida y sin miedo, y una segunda etapa que se caracteriza también por rapidez y falta de miedo, pero con más ideas que aparecen con mayor frecuencia y mayor claridad. Adicionalmente, como sostienen los mismos autores, cuanto más sepa el escritor, más difícil le resulta sintetizar todo su conocimiento y menos se preocupa por llenar páginas.

De lo anterior se desprende que la cantidad y la velocidad con que se producen las palabras, efectivamente, constituyen dimensiones de la fluidez escrita. No obstante, escribir mucho no significa necesariamente escribir bien. Como lo

ejemplifican Fellner y Apple (2006), si tal fuera el caso, un estudiante que utiliza repetidamente la misma oración simple dentro de un intervalo de tiempo determinado se consideraría más fluido aunque no lo fuera.

Estos autores definen la fluidez escrita como el número de palabras producidas en un plazo de tiempo determinado, considerando también la frecuencia léxica de estas palabras –aunque no la ortografía o el contenido–, siempre que el significado proporcionado por el escritor resulte fácilmente comprensible (véase Fellner & Apple, 2006:19). Partiendo de esta definición, es necesario hacer dos puntualizaciones. En primer lugar, si bien el uso de palabras sofisticadas puede tener una influencia positiva en la percepción de la fluidez por parte del potencial lector, tal y como se argumentó anteriormente, el vocabulario es un aspecto que debería abordarse de manera separada. En segundo lugar, se está de acuerdo con la observación de Fellner y Apple (2006) que la fluidez escrita debería estudiarse con independencia de la ortografía, aspecto que pertenece al área de la corrección, y del contenido, que alude a propiedades cualitativas del discurso y, probablemente, emerge de una combinación de otras dimensiones de la lengua.

De acuerdo con lo expuesto, se podría concebir la fluidez escrita como la habilidad del aprendiente de LE de acceder y recuperar segmentos de su memoria y traducir de manera eficiente su pensamiento dentro de un intervalo de tiempo determinado, tanto en términos de cantidad como de velocidad, y de modo equiparable, aunque no necesariamente idéntico, al de un hablante o escritor competente, nativo o no. Tal definición está en consonancia con los planteamientos de Wolfe-Quintero et al. (1998), quienes apuntan que aprendientes que cuentan con la misma cantidad de estructuras productivas pueden recuperarlas con diferentes grados de eficiencia, por tanto, la fluidez se puede medir a partir del número de estructuras incluidas en el discurso escrito dentro de un intervalo de tiempo fijo y bajo la asunción de que existe cierta homogeneidad en cuanto al nivel de dominio lingüístico del grupo meta.

Adicionalmente, sin desestimar la importancia de la precisión en la percepción de un discurso fluido, se puede entender la fluidez escrita como la integración de segmentos lingüísticos en oraciones que presentan cierto grado de precisión (véase Wolfe-Quintero et al., 1998). Por último, es necesario subrayar que la fluidez escrita no es una habilidad totalmente automática y libre de control atencional y esto concierne tanto a la LM como a la LE, aunque un alto grado de automatización de ciertos componentes de la escritura podría fomentarla (Olive & Kellogg, 2002; Fayol, Foulin, Maggio, & Lété, 2012).

En lo que se refiere a las medidas de fluidez escrita, siguiendo a Wolfe-Quintero et al. (1998), se pueden identificar tres tipos: (1) las medidas de frecuencia absoluta (número total de PAL, OR, CL, UT o verbos); (2) las que se basan en la velocidad de producción (por ejemplo, PAL/minuto); (3) las medidas de longitud (véase apartado 2.1) y las ratios que se basan en el número total de PAL en relación con el número total de UTLE y CLLE.

Un punto controvertido y destacado anteriormente concierne a las medidas de longitud. Según Wolfe-Quintero et al. (1998), estas medidas son propias de la

fluidez, ya que no evalúan las diferentes maneras en que se puede establecer la longitud de un texto. En cuanto a las medidas de longitud que se basan en las UTLE y CLLE (PAL/UTLE, PAL/CLLE), el supuesto que subyace a su uso es que en los niveles avanzados del aprendizaje de una LE los aprendientes son capaces de producir estructuras no solo más largas, sino también más correctas. Es por ello que Wolfe-Quintero et al. (1998) sostienen que estas medidas no reflejan el grado de corrección de una composición, sino más bien la fluidez alcanzada.

Otras objeciones respecto a las medidas de fluidez escrita versan sobre los siguientes puntos: (a) la extensión de un texto puede depender de factores como la familiaridad con el tema de la escritura o la decisión propia del escritor de incluir en su discurso una cantidad específica de palabras; (b) algunas medidas basadas en la frecuencia absoluta de las unidades de análisis UT, CL u OR parecen reflejar aspectos cualitativos de la composición escrita; (c) el uso de medidas de velocidad puede ser cuestionable, ya que es posible que el aprendiente no dedique mucho tiempo al desempeño de una tarea debido a una actitud negativa hacia la escritura; del mismo modo, un escritor competente puede que produzca pocas palabras no por ser menos fluido, sino por revisar su texto de manera más pormenorizada (Abdel Latif, 2008).

En síntesis, a diferencia del constructo de fluidez oral que ha sido bastante bien definido y ampliamente estudiado en la investigación empírica de ASL, la revisión de la bibliografía muestra la falta de precisión en la definición de la fluidez escrita y los consiguientes problemas metodológicos que supone su medición en el plano empírico. Las afirmaciones precedentes cobran aún mayor importancia si se considera el debate en torno a la idoneidad de algunas medidas de fluidez escrita, el grado en que estas reflejan dimensiones representativas del constructo, así como la variedad de factores que pueden incidir en la extensión de una composición escrita. Según parece, la consideración conjunta de estos factores y la aplicación de técnicas estadísticas como el AFE podrían conducir a la desambiguación del concepto y al establecimiento de algunas directrices metodológicas que permitan su medición consistente y sistemática.

5. Conclusiones

La complejidad, la precisión, la fluidez y el léxico son, sin lugar a dudas, constructos de carácter multidimensional y multifacético (Housen & Kuiken, 2009; Larsen-Freeman, 2009; Palloti, 2009; Housen et al., 2012) y de enorme utilidad para esclarecer cuestiones relacionadas con el procesamiento lingüístico y las prácticas educativas. Ahora bien, como se ha visto a lo largo de la revisión presentada en este artículo, respecto a la conceptualización de estos constructos y a la idoneidad de las medidas desarrolladas para su evaluación muchos interrogantes continúan vigentes.

A fin de obtener resultados fiables y también comparables, lo ideal sería utilizar varias medidas y conseguir de este modo superar una de las limitaciones metodológicas frecuentes en estudios de corte cuantitativo en ASL. No obstante, como bien señala Palloti (2009), aplicar todas las medidas propuestas en la

bibliografía especializada cada vez que se estudia la actuación y el desarrollo interlingüístico resulta especialmente difícil. Las dificultades no se deben solo a la cantidad de tiempo que requiere el tratamiento de los datos que arrojan las tareas de producción en una LE, sino también a la escasez de programas y herramientas específicamente diseñadas para el análisis de idiomas diferentes al inglés, así como a la cantidad de entrenamiento que supone su manejo.

En términos de análisis estadístico, la inclusión en un estudio empírico de medidas que evalúen las mismas facetas o facetas muy parecidas de un constructo concreto podría criticarse de ‘redundancia’ o crear problemas de singularidad y multicolinealidad al aplicarse ciertos métodos del AFE –como el Análisis Factorial Común– que, a pesar de su menor difusión en la investigación en Lingüística Aplicada, son extremadamente útiles para esclarecer cuestiones relacionadas con la representatividad de ciertas medidas en relación con determinados constructos.

Sumado a lo anterior, en las dimensiones lingüísticas de complejidad, precisión y fluidez no pueden obviarse las fluctuaciones atribuidas a una serie de factores como la edad, la maduración cognitiva y el nivel de educación formal alcanzado (Torrás & Celaya, 2001; Navés et al., 2003; Navés, 2007), capacidades cognitivas como la aptitud y la memoria operativa (Kormos & Trebits, 2011, 2012) o ciertas propiedades de las tareas implementadas.

Una última cuestión de especial relevancia concierne a la naturaleza de relación e interconexión entre estos constructos. Varios autores (Robinson, 2001, 2003; Skehan & Foster, 2001; VanPatten, 2002) han tratado de identificar aquellas áreas de la actuación que entran en competición [*trade-off*] por los recursos limitados del aprendiente de LE (por ejemplo, fluidez frente a complejidad, complejidad frente a precisión, precisión frente a fluidez, aunque puede haber más, por ejemplo, corrección lingüística frente a planificación o contenido) y es innegable el mérito de sus aportaciones. La postura que se mantiene en este artículo no es que no se produzcan *trade-offs* entre los constructos de complejidad, precisión y fluidez; indudablemente existen bajo ciertas condiciones de actuación. No obstante, parece haberse relegado a un segundo plano el reconocimiento de las diferencias individuales (cognitivas, de personalidad) como uno de los principales causantes de estos *trade-offs*. Preguntas como la siguiente: ¿existen diferencias en los patrones de *trade-off* entre aprendientes con alta y baja capacidad de memoria operativa, entre aprendientes con alta y baja motivación, extrovertidos e introvertidos?, todavía no cuentan con una respuesta satisfactoria (pero véase Kormos & Trebits, 2011).

Sea cual sea el caso, lo cierto es que los constructos de CAF interactúan de manera muy compleja y tal interacción podría reflejar diferentes caminos hacia el desarrollo del sistema de la IL por parte de diferentes aprendientes (Larsen-Freeman, 2006, 2009). Especialmente reveladoras son las siguientes palabras de Larsen-Freeman (2009):

In other words, the study of CAF has perhaps reached a point where the typical (reductionist) approach of taking factors one by one to see what effect each has on learner performance in a linear causal way does little to advance our understanding. This is because there is always some contingency, some limitation, and some outlier for even the most reliable of findings. This in itself might not be so worrisome -after all, it keeps us happily employed. However, there is the complication that as we explore CAF, as we look deeper into the three -complexity, accuracy, fluency- each becomes more complicated, with little end in sight. Furthermore, if we examine the dimensions one by one we miss their interaction and the fact that the way that they interact changes with time as well. (Larsen-Freeman, 2009:582)¹

Considerando conjuntamente lo expuesto, es necesario aproximarse a la complejidad, la precisión, la fluidez y el léxico como un conjunto dinámico, interrelacionado, variable, en constante cambio y dependiente del propio individuo (Larsen-Freeman, 2009; Housen et al., 2012). Los resultados son todavía demasiado provisionales y dejan abierto un amplio espectro de opciones para los investigadores que trabajan en el ámbito de ASL.

Referencias bibliográficas

- Abdel Latif, M. M. (2008), “*What do we mean by writing fluency? Proposing a new measure for assessing fluent written language production*”, paper presented at the 41st Annual Meeting of the British Association for Applied Linguistics, Swansea University, Wales, UK.
- Bruton, D. L. & D. L. Kirby (1987), “Research in the classroom: Written fluency: didn’t we do that last year?”, *The English Journal*, 76(7):89-92.
- Bulté, B. & A. Housen (2012), “Defining and operationalising L2 complexity”, en A. Housen, F. Kuiken & I. Vedder (eds.), *Dimensions of L2 performance and proficiency. Complexity, accuracy and fluency in SLA*. Amsterdam, The Netherlands: John Benjamins, 21-46.
- Chambers, F. (1997), “What do we mean by fluency?”, *System*, 25(4):535-544.
- Checa García, I. (2005), “Medidas de madurez sintáctica aplicadas a lecturas de ELE”, *Interlingüística*, 16(1):273-285.

¹ “En otras palabras, el estudio de los constructos de CAF quizás haya llegado a un punto en el que la aproximación típica (reduccionista) de considerar los factores uno por uno para ver qué efecto tiene cada uno de ellos en la actuación del aprendiente de una manera causal lineal permite avanzar poco en la comprensión del problema. Esto se debe a que siempre hay alguna contingencia, alguna limitación, y algunos valores atípicos, incluso para los resultados más fiables. Esto en sí mismo podría no ser tan preocupante -después de todo, nos mantiene felizmente ocupados. Sin embargo, existe la complicación de que a medida que exploramos los constructos de CAF, al mirar más profundamente en los tres -complejidad, precisión, fluidez- cada uno se vuelve más complicado, con pocas soluciones a la vista. Adicionalmente, si examinamos estas dimensiones una por una perdemos su interacción, así como el hecho de que la manera en que interactúan cambia también con el paso del tiempo” [Traducción propia].

- Cook, V. (1999), “Going beyond the native speaker in language teaching”, *TESOL Quarterly*, 33(2):185-209. doi:10.2307/3587717
- Crookes, G. (1990), “The utterance and other basic units for second language discourse analysis”, *Applied Linguistics*, 11(2):183-199. doi:10.1093/applin/11.2.183
- Ellis, N. C. & D. Larsen-Freeman (2006), “Language emergence: Implications for Applied Linguistics - Introduction to the Special Issue”, *Applied Linguistics*, 27(4):558-589. doi:10.1093/applin/aml028
- Fayol, M., Foulin, J. N., Maggio, S. & B. Lété (2012), “Towards a dynamic approach of how children and adults manage text production”, en E. L. Grigorenko, E. Mambrino & D. D. Preiss (eds.), *Writing: A mosaic of new perspectives*. New York: Taylor & Francis Group, LLC. 141-158.
- Fellner, T. & M. Apple (2006), “Developing writing fluency and lexical complexity with blogs”, *The JALT CALL Journal*, 2(1):15-26.
- Foster, P. & P. Tavakoli (2009), “Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity”, *Language Learning*, 59(4):866-896. doi:10.1111/j.1467-9922.2009.00528.x
- Foster, P., Tonkyn, A. & G. Wigglesworth (2000), “Measuring spoken language: A unit for all reasons”, *Applied Linguistics*, 21(3):354-375. doi:10.1093/applin/21.3.354
- Götz, S. (2013), *Fluency in native and nonnative English speech*. Amsterdam, The Netherlands: John Benjamins.
- Hirano, K. (1989), “Research on T-unit measures in ESL”, *Bull. Joetsu Univ. Educ.*, 8(2):67-77.
- Housen, A. & F. Kuiken (2009), “Complexity, accuracy, and fluency in Second Language Acquisition”, *Applied Linguistics*, 30(4):461-473. doi:10.1093/applin/amp048
- Housen, A., Kuiken, F. & I. Vedder (2012), “Complexity, accuracy and fluency: Definitions, measurement and research”, en A. Housen, F. Kuiken & I. Vedder (eds.), *Dimensions of L2 performance and proficiency. Complexity, accuracy and fluency in SLA*. Amsterdam, The Netherlands: John Benjamins, 1-20.
- Hunt, K. W. (1965), “Grammatical structures written at three grade levels”, NCTE Research Report No. 3. Champaign, IL: National Council of Teachers of English.
- Ishikawa, S. (1995), “Objective measurement of low-proficiency EFL narrative writing”, *Journal of Second Language Writing*, 4(1):51-69. doi:10.1016/1060-3743(95)90023-3
- Iwashita, N., Brown, A., McNamara, T. & S. O’Hagan (2008), “Assessed levels of second language speaking proficiency: How distinct?”, *Applied Linguistics*, 29(1):24-49. doi:10.1093/applin/amm017
- Johansson, V. (2008), “Lexical diversity and lexical density in speech and writing: A developmental perspective”, *Lund Working Papers in Linguistics*, 53:61-79.

- Koizumi, R. (2012), “Relationships between text length and lexical diversity measures: Can we use short texts of less than 100 tokens?”, *Vocabulary Learning and Instruction*, 1(1):60-69. doi:10.7820/vli.v01.1.koizumi
- Kormos, J. (2011), “Task complexity and linguistic and discourse features of narrative writing performance”, *Journal of Second Language Writing*, 20(2):148-161. doi:10.1016/j.jslw.2011.02.001
- Kormos, J. & M. Dénes (2004), “Exploring measures and perceptions of fluency in the speech of second language learners”, *System*, 32:146-164.
- Kormos, J. & A. Trebits (2011), “Working memory capacity and narrative task performance”, en P. Robinson (ed.), *Second language task complexity. Researching the Cognition Hypothesis of language learning and performance*. Amsterdam, The Netherlands: John Benjamins, 267-285.
- Kormos, J. & A. Trebits (2012), “The role of task complexity, modality, and aptitude in narrative task performance”, *Language Learning*, 62(2):439-472. doi:10.1111/j.1467-9922.2012.00695.x
- Larsen-Freeman, D. (2006), “The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English”, *Applied Linguistics*, 27(4):590-619. doi:10.1093/applin/aml029
- Larsen-Freeman, D. (2009), “Adjusting expectations: The study of complexity, accuracy, and fluency in Second Language Acquisition”, *Applied Linguistics*, 30(4):579-589. doi:10.1093/applin/amp043
- Laufer, B. & P. Nation (1995), “Vocabulary size and use: Lexical richness in L2 written production”, *Applied Linguistics*, 16(3):307-322. doi:10.1093/applin/16.3.307
- Lu, X. (2012), “The Relationship of lexical richness to the quality of ESL learners’ oral narratives”, *The Modern Languages Journal*, 96(2):190-208. doi:10.1111/j.1540-4781.2011.01232.x
- MacWhinney, B. (2000), *The CHILDES Project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- McCarthy, P. M. & S. Jarvis (2007), “Vocd: A theoretical and empirical evaluation”, *Language Teaching*, 24(4):459-488. doi:10.1177/0265532207080767
- McCarthy, P. M. & S. Jarvis (2010), “MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment”, *Behavior Research Methods*, 42(2):381-392. doi:10.3758/BRM.42.2.381
- Meara, P. & H. Bell (2001), “P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts”, *Prospect*, 16(3):323-337.
- Navés, T. (2007), “Analytical measures of learners’ written interlanguage”, ponencia presentada en el XXV Congreso Internacional de AESLA, Universidad de Murcia, España.
- Navés, T., Torras, M. R. & M. L. Celaya (2003), “Long-term effects of an earlier start”, en S. Foster-Cohen & S. Pekarek (eds.), *EUROSLA-Yearbook. Annual Conference of the European Second Language Association*. Amsterdam, The Netherlands: John Benjamins, 103-130.

- Nariman-Jahan, R. & M. Rahimpour (2011), “The effects of planning and proficiency on language production of writing task performance”, *Educational Research*, 2(9):1528-1537.
- Olive, T. & R. T. Kellogg (2002), “Concurrent activation of high- and low-level production processes in written composition”, *Memory and Cognition*, 30(4):594-600. doi:10.3758/BF03194960
- Ortega, L. (2003), “Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing”, *Applied Linguistics*, 24(4):492-518. doi:10.1093/applin/24.4.492
- Palloti, G. (2009), “CAF: defining, refining and differentiating constructs”, *Applied Linguistics*, 30(4):590-601. doi:10.1093/applin/amp045
- Polio, C. G. (1997), “Measures of linguistic accuracy in second language writing research”, *Language Learning*, 47(1):101-143. doi:10.1111/0023-8333.31997003
- Richards, B. (1987), “Type/token ratios: what do they really tell us?”, *Journal of Child Language*, 14(2):201-209. doi:10.1017/S0305000900012885
- Richards, B. & D. Malvern (2000), “*Measuring vocabulary richness in teenage learners of French*”, paper presented at the British Educational Research Association Conference, Cardiff University.
- Robinson, P. (2001), “Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA”, en P. Robinson (ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 287-318.
- Robinson, P. (2003), “The Cognition Hypothesis, task design, and adult task-based language learning”, *Second Language Studies*, 21(2):45-105.
- Robinson, P., Cadierno, T. & Y. Shirai (2009), “Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production”, *Applied Linguistics*, 30(4):533-554. doi:10.1093/applin/amp046
- Schmidt, R. W. (1992), “Psychological mechanisms underlying second language fluency”, *Studies in Second Language Acquisition*, 14:357-385. doi:10.1017/s0272263100011189
- Segalowitz, N. (2007), “Access fluidity, attention control, and the acquisition of fluency in a second language”, *TESOL Quarterly*, 41(1):181-186. doi:10.1002/j.1545-7249.2007.tb00047.x
- Šišková, Z. (2012), “Lexical richness in EFL students’ narratives”, *University of Reading Language Studies Working Papers*, 4:26-36.
- Skehan, P. (1996), “A framework for the implementation of task-based instruction”, *Applied Linguistics*, 17(1):38-62. doi:10.1093/applin/17.1.38
- Skehan, P. (2003), “Task-based instruction”, *Language Teaching*, 36(1):1-14. doi:10.1017/S026144480200188X
- Skehan, P. (2009), “Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis”, *Applied Linguistics*, 30(4):510-532. doi:10.1093/applin/amp047

- Skehan, P. & P. Foster (1997), “Task type and task processing conditions as influences on foreign language performance”, *Language Teaching Research*, 1(3):185-211. doi:10.1177/136216889700100302
- Skehan, P. & P. Foster (2001), “Cognition and tasks”, en P. Robinson (ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 181-205.
- Sotillo, S. M. (2000), “Discourse functions and syntactic complexity in synchronous and asynchronous communication”, *Language Learning & Technology*, 4(1):82-119.
- Torras, M. R. & M. L. Celaya (2001), “Age-related differences in the development of written production. An empirical study of EFL school learners”, *International Journal of English Studies*, 1(2):130-126.
- Treffers-Daller, J. (2013), “Measuring lexical diversity among L2 learners of French: an exploration of the validity of D, MTL D and HD-D as measures of language ability”, en S. Jarvis & M. Daller (eds.), *Vocabulary knowledge: Human ratings and automated measures*. Amsterdam, The Netherlands: John Benjamins, 79-104.
- Unsworth, S. (2008), “Comparing child L2 development with adult L2 development: How to measure L2 proficiency”, en B. Haznedar & E. Gavrusena (eds.), *Current trends in child second language acquisition: A generative perspective*. Amsterdam, The Netherlands: John Benjamins, 301-336.
- Vajjala, S. & D. Meurers (2012), “On improving the accuracy of readability classification using insights from second language acquisition”, paper presented at the 7th Workshop on Innovative Use of NLP for Building Educational Applications, The Association for Computational Linguistics, Montréal, Canada.
- VanPatten, B. (2002), “Processing Instruction: An update”, *Language Learning*, 52(4):755-803.
- Woerfel, T. & S. Yilmaz (2011), “Lexical development of German-Turkish bilinguals: A comparative study in written discourse”, en C. Cummins et al. (eds.), *Proceedings of the 6th Cambridge Postgraduate Conference in Language Research*. Cambridge: Cambridge Institute of Language Research, 240-251.
- Wolfe-Quintero, K., Inagaki, S. & H. Y. Kim (1998), *Second language development in writing: Measures of fluency, accuracy & complexity* (Technical Report 17). Honolulu, HI: University of Hawai’i, Second Language Teaching and Curriculum Center.
- Yang, W. & Y. Sun (2015), “Dynamic development of complexity, accuracy and fluency in multilingual learners’ L1, L2 and L3 writing”, *Theory and Practice in Language Studies*, 5(2):298-308. doi:10.17507/tpls.0502.09
- Yu, G. (2007), “Lexical diversity of MELAB speaking and writing performances”, *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 5:79-116.