

Doktorsavhandlingen *El uso de combinaciones de palabras con que en un corpus de aprendices suecos de español como lengua extranjera* är ett resultat av ett samarbetsavtal mellan Göteborgs universitet och universitetet i Barcelona (Universitat de Barcelona), och belyser hur typiska formuleringar i svenska språket påverkar användningen av det spanska ordet *que* (som på svenska motsvarar flera olika ord, såsom *att/som/vilket/vilken/vilka/än/vad*). Materialet som använts utgörs av skrivuppgifter producerade av 45 svensktalande studenter som under 2008-2010 läste spanska vid Göteborgs universitet och vid Högskolan i Jönköping. Studien använder sig av korpuslingvistik och utgörs av en kontrastiv analys av det s.k. interimspråket, dvs. ett mellanspråk som under inläringen skapas av varje inlärare och utgörs av en blandning av det egna och det främmande språket under inläringen. Avhandlingen pekar på hur dessa inlärningsstrategier ibland resulterar i positiva kommunikativa resurser (det är kanske inte så spanskinfödda säger, men det blir begripligt) men ibland blir det i mindre lyckade användningar (som strider mot målspråkets grammatik och syntax).

AYMÉ PINO RODRÍGUEZ

El uso de combinaciones de palabras con *que* en un corpus de aprendices suecos de español como lengua extranjera

El uso de combinaciones de palabras con *que* en un corpus de aprendices suecos de español como lengua extranjera

N Concordance

1 hombre, Edward Lewis (Richard Gere) **que** trabaja como un hombre de
 2 rico y afortunado. Pero hay un **Algo que es** en su vida, el amor. Edward va
 3 Roberts) para ser su profesor. **Algo que es** no cuenta con es que el va
 4 Algo que Edward no **con es que** el va a enamorarse con ella, una mujer
 5 Pretty Woman **una película que se trata** de muchos temas diferentes y
 6 son muy diferentes pero **hay algo que los maneja** unido; la soledad y el
 7 famosas y compra ropa. **Yo creo que esta** es una película para la soledad, a hacer compras
 8 de todas las chicas. Es **una película que se dice** para los espectadores jóvenes,
 9 primero a chicas pero también a chicos que **se dice** soñarse a un mundo donde
 10 pero me gusta viajar y hay **varias que se dice** que debería visitar. Pues bien,
 11 iríamos a Creta en Grecia. Me recuerdo **que se dice** que empezó muy temprano el día del
 12 una piscina grande. Durante la semana **que se dice** que en Creta bañamos mucho
 13 el avión estuvo atrasado y tuvimos **que se dice** que en Creta bañamos mucho
 14 buen tiempo" y se trata de una familia **que se dice** que se va a vacaciones con su
 15 no sabemos adonde van. "Tenemos **que se dice** que apura a decir él. La mamá trata
 16 solo en el gimnasio también. Creo **que es más** divertido entrenar en grupo
 17 es más divertido entrenar en grupo **ya que se puede** estar juntos y motivar
 18 y muchos amigos. Algo otro divertido **que a veces he** pasado mi tiempo libre es
 19 estoy buena de jugar al tenis pero creo **que es un deporte** muy divertido.
 20 ahora, solía bailar el rock and roll. Creo **que es muy divertido** bailar el rock and
 21 si la música es rápida. Uno tiene **que tener** condición física para tener
 22 solo en el gimnasio también. Creo **que es más divertido** bailar en grupo
 23 es más divertido entrenar en grupo **ya que se puedo** luchar juntos y motivar
 24 y muchos amigos. Algo otro divertido **que a veces hago** en mi tiempo libre es
 25 estoy buena de jugar al tenis pero creo **que es un deporte** muy divertido.
 26 ahora, solía bailar el rock and roll. Creo **que es muy divertido** bailar el rock and
 27 si la música es rápida. Uno tiene **que tener** condición física para tener
 28 de la vida y de la cultura de los países **que he** conocido. Mi primer destino era
 29 a estudiar en una escuela privada **que se llama** ECELA. Estudiaba por dos
 30 super divertido! Eramos cinco personas **que celebramos** juntos, dos suecas,
 31 hasta Bolivia. Fuimos a una ciudad **que se llama** Copacabana, que está a
 32 una ciudad que se llama Copacabana, **que está** a orillas del lago Titicaca. Eran
 33 lago y también visitamos Isla del sol, **que es** una isla en el lago, donde, según
 34 Había mucha pobreza y muchos niños **que trabajaban** por las calles y esto nos
 35 Argentina y Buenos Aires, una ciudad **que a mí me encanta!** Hay todo en esta
 36 me gustó el país mucho y mi plan es **que vuelva** de nuevo para vivir allá por tal
 37 a Suecia. Eran unos meses increíbles **que nunca voy** a olvidar.



GÖTEBORGS UNIVERSITET
 INST FÖR SPRÅK OCH LITTERATURER

Aymé Pino Rodríguez

Departamento de lenguas y literaturas
 Universidad de Gotemburgo

ISBN: 978-91-628-8488-8

El uso de combinaciones de palabras con *que* en un corpus de aprendices suecos de español como lengua extranjera

Aymé Pino Rodríguez



GÖTEBORGS UNIVERSITET
INST FÖR SPRÅK OCH LITTERATURER

Avhandling för filosofie doktorexamen i spanska
Göteborgs universitet, 2012-06-02

Disputationsupplaga

© Aymé Pino Rodríguez, 2012

Tryck: Reprocentralen, Humanistiska fakulteten, Göteborgs universitet

Omslag: Thomas Ekholm

ISBN: 978-91-628-8488-8

<http://hdl.handle.net/2077/28835>

A los que no los detienen las sillas del camino

ABSTRACTS

Title: *El uso de combinaciones de palabras con que en un corpus de aprendices suecos de español como lengua extranjera*/The Use of Word Combinations with *que* in a Corpus of Swedish Learners of Spanish as a Foreign Language.

Previous research in phraseological studies has approached the difficulties that learners of foreign and second languages have to produce word combinations derived from idiomatic- or phraseological principles. Using the theoretical framework of Sinclair's idiom principle and applying a corpus-based methodological approach to contrastive linguistics and the study of interlanguage, I have analyzed the use of Spanish word sequences by Swedish learners. I argue that Swedish learners of Spanish as a Foreign Language overuse different word combinations due to mother tongue influence (Swedish), and as a consequence of different accommodation strategies in their interlanguage. The differences in function and frequency of use of these combinations decrease proportionally with the development of learners' knowledge and skills in the target language.

The thesis presents a quantitative and qualitative analysis of word combinations with *que*: *lo que*, *de que*, *algo que*, *dice que* in 135 texts (corpus SAELE-Swedish students of Spanish as a foreign language) produced by 45 Swedish learners of Spanish as a Foreign Language (27 women and 18 men), at two Swedish universities. The sample represents 25.4% ($P < 0.005$) of the total of 177 students (universe) enrolled in academic courses in Spanish as a Foreign Language. The study is contrastive corpus analysis of argumentative texts written by intermediate-level Swedish learners of Spanish (Corpus SAELE compiled by the author) which was compared with two control groups: one of native speakers of Spanish (control corpus of Spanish CEDEL2) and another of native speakers of Swedish (control corpus of Swedish ARGUS).

The main research question was to examine whether Swedish learners of Spanish as a Foreign Language tend to use different word combinations with *que*: (*lo que*, *de que*, *algo que* and *dice que*) than Spanish and Swedish-natives. In addition, I have aimed at: (a) to carry out a survey based on sociolinguistic data obtained from students and b) to investigate if the level of proficiency of Swedish learners of Spanish Foreign Language corresponded to A2-B1 within Common European Framework of Reference for Languages.

Due to its polysemous nature, *que* has a high frequency of use in the Spanish control corpora (3.45%). The Swedish equivalents *att* appeared with a frequency of 3.75% and *som* with of 2.13% in the Swedish control corpus. In the corpus SAELE *que* appeared with a frequency of 2.95%.

The study has shown the differences in use and frequency in the production of Swedish learners with Spanish and Swedish speakers. The combination *dice que* seems to be the most frequently used by Swedish learners of Spanish as a foreign language, although it was 21.6 % more frequent among the Swedish learners than Spanish-speakers (control corpus CEDEL2) and 13.44 % than among Swedish-speakers (control corpus ARGUS). The overuse of the combination *dice que* could reflect a tendency by Swedish learners of Spanish to use formulations typical of speech rather than of writing. In addition it was found that Swedish learners of Spanish use the combination *algo que* (0.4%) double as often as Spanish-speakers and *lo que* equivalent to *det som* 6.57% and to *vad som* 7.47 % more frequently than Swedish-speakers. It shows learners linguistics *continuum* if you consider that *lo que* is the combination most frequently used by Spanish-speakers and the Swedish equivalents to *algo que* were the combinations most commonly used by Swedish-speakers.

The analyses of the essays by the Swedish learners of Spanish as a Foreign Language participating in the present study showed that their skills conformed to level A2 to B1 (Common European Framework of Reference for Languages) according to the analyses of their essays. The learners produced some word combinations similar to those produced by Spanish natives, but made some errors in their written production. It was found that individual differences among participants with regard to age, length of study of Spanish in Sweden or a Spanish-speaking country could be another factor to take in account. A comparison of the Swedish learners with more advanced students of Spanish (level B1-B2), indicated that the latter used particle in a way that is directly proportional to the preferred choice by a native speaker. Pedagogical conclusions will be used to develop concordance and collocation-based exercises as a way of raising students' awareness of the use and function of recurrent words combinations with *que* in their written production.

Keywords: Corpus linguistics, Contrastive analysis, Learners corpus, Spanish as a Foreign Language, Word combinations, Idiom principle, Collocations, Interlanguage.

AGRADECIMIENTOS

Durante el trayecto de realización de esta tesis me han tendido la mano muchas personas en diferentes países, lugares y contextos. A todas ellas vaya mi eterna gratitud.

En especial quisiera agradecer a mis tutores por su sabiduría, generosidad, ejemplo de laboriosidad y constancia, tolerancia e incondicional apoyo.

A Ken Benson, por descubrir cosas entre líneas en sus revisiones, su rigurosidad metodológica que de por sí es una enseñanza para un doctorando, su sistemático apoyo práctico y sus palabras de aliento cuando las necesitaba.

A Isabel Verdager Clavera porque guió mis primeros pasos en la lingüística de corpus desde el primer seminario de doctorado, siguió mis empeños en el trabajo del DEA y continuó creyendo en mi idea de hacer un corpus de aprendices suecos de E/LE.

A la Universidad de Gotemburgo y la fundación Bo Linderöth-Olssons con cuyo apoyo financiero pude realizar una gran parte de mi doctorado, asistir a diversas conferencias internacionales y realizar una estancia investigadora en la Universitat de Barcelona.

A los compañeros del departamento de lenguas y literaturas que de una u otra manera me han ayudado en mi período de estudio y trabajo, en especial a Nicole por su apoyo moral y administrativo. A los profesores de los cursos de nivel avanzado en que participé: Ingmar Söhrman, Andrea Castro y Alejandro Urrutia. A todos los colegas de la sección de español que participaron en los seminarios de doctorado en que presenté mis textos por sus críticas constructivas, sugerencias y preocupación.

A la Universitat de Barcelona por acogerme como estudiante de doctorado y donde siempre me siento como en casa. En particular a la facultad de filología y el departamento de inglés y alemán.

A mis profesores Ramón Ribé, Margarida Cambra, Elsa Tragant y María Luz Celaya quienes me abrieron las puertas al mundo de la lingüística aplicada y a Mar Cruz Piñol (del departamento de Hispánicas) por su lectura crítica en el seminario final.

A mis profesores de la universidad de Linköping Ingrid Hermerén, Monica Sandlund y Anna Forné, a quienes recuerdo por sus enseñanzas en el tiempo que estudié allí para obtener mi diploma sueco de profesora de español e historia.

A Lucía por la revisión de los capítulos, a Linda por ayudarme con el formato y a Thomas con la impresión y el diseño de la portada. A Daniel por la instalación del procesador de texto y los programas y a Pernilla y Oriol por enseñarme algunas de sus sutilezas.

A mis compañeros del instituto superior de Jönköping por su apoyo durante estos años, especialmente a la sección de lenguas y a Catarina, Bertil, Jörgen, Mari-Ann, Maj, Kerstin y Ann. A Teresita, Lucía, Libertad y Birgit por hacerse responsable de los cursos de español durante mi ausencia.

A mis hijos Luis Alejandro, Ana y Camila por cumplir siempre con sus deberes y ser tolerantes ante mi desasosiego durante muchos días en que no me salía la escritura como quería, a Cami por ayudarme con los diagramas y cuadros; a mi esposo por su amor solidario, sus fines de semana de revisiones de la tesis y sus miles de consejos de cómo mejorar uno u otro aspecto del trabajo que acogió como si fuese propio y a Luni por su fidelidad y los paseos por el bosque.

Por último, pero no por ello menos importante quisiera agradecer a todos mis alumnos por su ingenio en encontrar estrategias para aprender español, y por ser la razón principal de mi trabajo y de este estudio.

ÍNDICE

ABREVIATURAS	13
CAPÍTULO 1. INTRODUCCIÓN.....	14
1.1 Objetivos	17
1.2 Esquema metodológico	19
1.3 Disposición.....	20
CAPÍTULO 2. PANORAMA DEL DESARROLLO DE LA LINGÜÍSTICA DE CORPUS Y LOS CORPUS DE APRENDICES	23
2.1 La lingüística de corpus como herramienta metodológica.....	23
2.1.2 Análisis histórico del desarrollo de la lingüística de corpus	27
2.1.3 Tradición e investigaciones en lingüística de corpus en Escandinavia y en Suecia en particular.....	32
2.2 Aplicación pedagógica de la LC y los corpus de aprendices	35
CAPÍTULO 3. ANÁLISIS CONTRASTIVO Y ENSEÑANZA-APRENDIZAJE DE LENGUAS	40
3.1 La lengua materna, la lengua extranjera y la interlengua del aprendiz de idioma	40
3.2 El análisis contrastivo de la interlengua	42
3.3 El proceso de enseñanza-aprendizaje de E/LE en la universidad de Gotemburgo y en el Instituto superior de Jönköping en Suecia.....	47
3.3.1 El contexto situacional del estudio.....	48
3.4 Las destrezas según el MCERL	51
3.4.1 La destreza escrita y las competencias comunicativas.....	54
CAPÍTULO 4. COMBINACIONES DE PALABRAS CON <i>QUE</i>	60
4.1 Combinaciones de palabras	61
4.2 Estudios precedentes de combinaciones de palabras con <i>que</i> entre nativos suecos aprendices de E/LE.....	64
4.3 Combinaciones de palabras con <i>que</i>	65
4.3.1 Combinación <i>lo que</i>	72
4.3.2 Combinación <i>de que</i>	74
4.3.3 Combinación <i>algo que</i>	77
4.3.4 Combinación <i>dice que</i>	78
CAPÍTULO 5. MATERIAL DE TRABAJO	81
5.1 El corpus SAELE	81
5.1.1 Metodología	82
5.2 Los corpus de control: ARGUS y CEDEL2.....	88
5.2.1 El corpus ARGUS	88
5.2.2 El corpus CEDEL2.....	89
5.3 La encuesta y sus parámetros metodológicos	90
5.3.1 Los participantes en el estudio	93
CAPÍTULO 6. RESULTADOS CUANTITATIVOS DE LOS CORPUS	99
6.1 Resultados del corpus SAELE	99
6.2 Resultados del corpus de control CEDEL2.....	110
6.3 Resultados del corpus de control ARGUS	118
6.4 Resultados de la comparación de frecuencias entre el corpus SAELE y el subcorpus de nativos hispanohablantes del CEDEL2	131
CAPÍTULO 7. DISCUSIÓN DE LOS RESULTADOS	135
7.1 Uso de las combinaciones de palabras con <i>que</i> : <i>lo que</i> , <i>de que</i> , <i>algo que</i> y <i>dice que</i> por los aprendices suecos de E/LE	135
7.1.1 Uso de la combinación <i>lo que</i> por los aprendices suecos de E/LE	141

7.1.2	Uso de la combinación <i>de que</i> por los aprendices suecos de E/LE.....	143
7.1.3	Uso de la combinación <i>algo que</i> por los aprendices suecos de E/LE	146
7.1.4	Uso de la combinación <i>dice que</i> por los aprendices suecos de E/LE.....	147
7.1.5	Correspondencia entre el uso de las combinaciones de palabras con <i>que: lo que, de que, algo que y dice que</i> entre aprendices suecos de E/LE, nativos hispanohablantes y nativos suecos.....	149
7.1.6	La influencia de la lengua materna	155
7.2	Caracterización sociolingüística del aprendiz sueco de E/LE participante del corpus SAELE	157
7.3	Correspondencia entre los descriptores del MCERL y la competencia de los aprendices suecos de E/LE participantes en el estudio	162
7.3.1	Perfil del aprendiz sueco de E/LE participante en el estudio	166
CAPÍTULO 8. CONCLUSIONES.....		168
BIBLIOGRAFÍA.....		174
LISTA DE ANEXOS		198
Anexo 1. Plan de estudio del <i>Grundkurs 1-30hp</i> , cursos 2008/09 y 2009/10. Universidad de Gotemburgo.....		199
Anexo 2. Plan de estudio curso <i>Spanska 1-30hp</i> , cursos 2008/09 y 2009/10. Instituto superior de Jönköping.		203
Anexo 3. Carta de presentación de la tesis a los participantes.....		208
Anexo 4. Certificado de consentimiento informado.		210
Anexo 5. La encuesta		211

ABREVIATURAS

AC: Análisis contrastivo

ACI: Análisis contrastivo de la interlengua

AELINCO: Asociación española de lingüística de corpus

ARGUS: Corpus de aprendices (nativos suecos- corpus de control en esta tesis)

CEDEL2: Corpus de aprendices (inglés L1, español L2; español L1 – el subcorpus de nativos hispanohablantes sirve de corpus de control en esta tesis)

E/LE: Español como lengua extranjera

HSV: Högskoleverket. Dirección nacional de enseñanza superior sueca

IL: Interlengua

LA: Lingüística aplicada

LE: Lengua extranjera

LC: Lingüística de corpus

L1: Lengua primera

L2: Lengua segunda

LM: Lengua materna

MCERL: Marco común europeo de referencia para las lenguas

PEL: Portfolio europeo de lenguas

SAELE: Corpus de suecos aprendices de español como lengua extranjera

CAPÍTULO 1. INTRODUCCIÓN

La tesis *El uso de combinaciones de palabras con que en un corpus de aprendices suecos de español como lengua extranjera* se ha realizado en el marco de un convenio de colaboración entre la universidad de Gotemburgo y la Universitat de Barcelona y se inscribe, de acuerdo con su contenido, en el área de la lingüística aplicada (en adelante, LA) centrándose en un aspecto de la destreza escrita y competencia comunicativa en el aprendizaje del español como lengua extranjera (en adelante, E/LE).

Según un informe de la Dirección nacional de enseñanza superior sueca¹ los estudios sobre adquisición y enseñanza-aprendizaje de lenguas han crecido sostenidamente en las últimas décadas en diferentes niveles educativos, incrementándose el número de publicaciones de tesis doctorales que abordan cuestiones relacionadas con esta área en los departamentos universitarios de francés e inglés (Enkvist, 2005: 84-85). Sin embargo, no existe la misma tradición de investigaciones en la asignatura E/LE, aunque en Suecia el español sea el idioma más frecuentemente elegido en la enseñanza secundaria y de bachillerato, la primera materia optativa de lenguas modernas después del inglés, en los últimos 10 años, (Österberg, 2008; Ferm & Malmberg, 2001), y represente incluso el idioma mayoritario entre las lenguas romances en el nivel universitario (Álvarez & Albanesi, 2006; Hejzlar et al., 2005).

Entre 2001-2010 se efectuaron 25 tesis doctorales en las universidades suecas sobre la lengua y literatura española², de ellas solamente una aborda un tema relativo a la enseñanza-aprendizaje, en el que específicamente su autora correlaciona la motivación y el desarrollo del español L2 en una muestra de estudiantes de E/LE (Österberg, 2008).

En el momento de inicio de la presente investigación existía una ausencia casi total de trabajos publicados cuyo objeto de estudio haya sido la lengua escrita de sujetos aprendices suecos de E/LE, lo cual fue uno de los acicates que nos motivó a contribuir a rellenar un vacío en esta área investigativa.

¹ Informe sobre estudios de lenguas en la universidad sueca. Véase en: http://gupea.ub.gu.se/bitstream/2077/18161/1/gupea_2077_18161_1.pdf [Fecha de consulta: 22-11-2011].

² Se hizo una búsqueda entre 33919 tesis doctorales registradas en: http://www.avhandlingar.se/om/spanska/?avancerad=ja&startrecord=&sok_title=&sok_creator=&sok_subject=spanska&sok_date=&sok_language=spa&sok_type=&sok_set=&sok_publisher=&fulltext= [Fecha de consulta: 22-01-2011].

Aparte de lo anterior, hubo además otros tres factores determinantes en la elección del objeto de estudio y la metodología aplicada. En primer lugar se debe apuntar que, en estudios precedentes se ha constatado que en la producción oral y escrita de los hablantes nativos hay una alternancia entre el principio de elección libre (*open-choice principle*), cuando hay libertad en la selección de las palabras, y el principio idiomático o la tendencia fraseológica del lenguaje (*idiom principle*) (Erman & Warren, 2000: 50-51). Sin embargo, parece ser que los aprendices de lenguas extranjeras (en adelante, LE) y de lenguas segundas (en adelante, L2) tienen más dificultades cuando utilizan determinadas combinaciones de palabras formadas a partir del principio idiomático o tendencia fraseológica del lenguaje (Flowerdew, 1998, 2003; Granger & Paquot, 2005; Paquot, 2005).

Estas asociaciones formadas por secuencias de dos o más palabras en forma continua o discontinua constituyen una unidad que relaciona significado y función en un contexto determinado, su elección refleja tanto la intuición como el ‘conocimiento compartido’³ del hablante (Moon, 1997: 44) y han recibido por los investigadores más de cincuenta denominaciones diferentes (Wray, 2002: 9).

En este trabajo se analiza el uso por aprendices suecos de E/LE (en comparación con otros nativos suecos y nativos hispanohablantes) de cuatro combinaciones idiomáticas de palabras con *que*: *lo que*, *de que*, *algo que*, *dice que*⁴ debido a que es precisamente con las combinaciones formadas a partir del principio idiomático donde los aprendices de lenguas extranjeras tienen más dificultades.

Esta elección resulta doblemente motivadora si se analiza cómo los nativos suecos aprendices de E/LE eligen *que* y sus combinaciones en la producción escrita de la lengua meta por varias razones. Por una parte, esta partícula en lengua española se corresponde, desde el punto de vista funcional, a por lo menos siete partículas diferentes en lengua sueca (*att*, *som*, *vilket/vilken/vilka*, *än* y *vad*), mientras que históricamente en español la conjunción y relativo *que* han monopolizado funciones muy diversas alcanzando un valor polisémico (Penny, 1993: 223). Por otra parte, porque el significado léxico y gramatical de una palabra o una unidad pluriverbal suele cambiar en dependencia del contexto de uso (Römer & Schulze, 2009: 2), y en

³ ‘Shared knowledge’.

⁴ En el Cap. 4 nos referimos a estudios precedentes de combinaciones de palabras, entre los cuales hay dos sobre las combinaciones de palabras con *que* (Pino, 2006, 2010).

INTRODUCCIÓN

dependencia de la forma en que sea utilizada, incluso en cohesión con otras palabras en el contexto lingüístico, va a incidir en que el mensaje del escritor llegue correctamente o no al receptor. Para que un aprendiz sueco de E/LE utilice las combinaciones de palabras con *que* de forma idiomática, se precisan conocimientos y habilidades que le permitan ‘reconocerlas’ por haberlas utilizado anteriormente y ‘reconstruirlas’ en asociación con otras palabras en la producción escrita u oral⁵.

En segundo lugar, sería interesante conocer si la producción escrita del aprendiz sueco de E/LE se corresponde con la descripción hecha en el MCERL respecto al nivel de destreza-competencia lingüística a partir de las habilidades mostradas en sus textos, aunque de acuerdo a nuestro propósito el estudio se limita al análisis del uso en textos elaborados por una muestra específica de estudiantes de E/LE de las combinaciones con *que: lo que, de que, algo que, dice que* y a la aplicación de una encuesta con que pretendemos identificar la posible existencia de otros factores sociolingüísticos que pueden haber incidido en los participantes del estudio (Cap.5).

En tercer lugar, nos pareció conveniente estudiar el uso de las combinaciones de palabras por aprendices suecos para tratar de implementar en el futuro actividades didácticas orientadas a ayudar a los nativos suecos a comprender qué aspectos de la lengua materna pueden interferir en el uso de esas combinaciones en la producción oral y escrita (Söhrman, 2007: 56; Ringbom, 2007) y el rol de algunas de las colocaciones más típicas (Herbst, 2009: 51), específicamente tratando de identificar posibles correspondencias con resultados de estudios anteriormente publicados que indican que aunque los aprendices incluso de nivel avanzado pueden producir una gran cantidad de colocaciones apropiadas, su intuición en la elección y la fluidez no siempre coincide con la de los nativos (Siyanova & Schmitt, 2008: 429).

Respecto a la metodología decidimos combinar dos herramientas: la lingüística de corpus (en adelante, LC) y el análisis contrastivo (en adelante, AC), lo que en la práctica resulta en un enfoque de análisis contrastivo de la interlengua⁶ (en adelante, ACI) del aprendiz sueco de E/LE basado en corpus (Tognini-Bonelli, 2002; Granger, 1996). Lo novedoso reside en que adaptamos estas metodologías para

⁵ De acuerdo con Hoey: *Every word is primed for use in discourse as a result of the cumulative effects of an individual's encounters with the word. If one of the effects of the initial priming is that regular word sequences are constructed, these are also in turn primed* (2005:13).

⁶ La interlengua es el sistema lingüístico empleado por un aprendiz de una LE o de una L2 que se encuentra entre el sistema de la lengua nativa (LN) y el de la lengua meta. En el capítulo 3 se desarrolla este aspecto.

investigar la interlengua de aprendices suecos de E/LE a partir de que otros investigadores ya la habían aplicado en sus estudios para explorar la lengua hablada y escrita de hablantes nativos, de L2 y LE con diferentes propósitos (Granger, 1998; Johansson, 2007, Siepman, 2005).

Esta distinción metodológica junto a los objetivos que nos trazamos se corresponde asimismo con la necesidad actual de crear corpus de aprendices con datos de diferentes registros de la lengua e informantes en diferentes niveles de aprendizaje cuya lengua materna sea diferente (Baralo, 2010) para conocer mejor los rasgos que caracterizan el *continuum* interlingüístico en que se integran la lengua materna y la lengua extranjera (Baralo, 2004; Cook, 2002). Fue así que a través del presente estudio compilamos un corpus de textos elaborados por aprendices suecos de E/LE que cursaban estudios en dos universidades suecas durante los cursos lectivos 2008/09 y 2009/10 (Véase el Cap.5).

Explicada la relevancia de la presente investigación y la motivación por la cual la hemos emprendido, se presentan en los apartados siguientes los objetivos e hipótesis (1.1), la metodología seguida (1.2) y aspectos sobre la disposición en la presentación de los resultados del presente trabajo (1.3).

1.1 Objetivos

El objetivo general es comparar el uso de algunas combinaciones de palabras de la producción escrita de aprendices adultos suecos de E/LE (nivel A2-B1 de acuerdo a la escala del MCERL⁷) con las producidas por nativos hispanohablantes y suecos. Este objetivo general dado su amplitud lo delimitamos en cinco objetivos específicos:

(1) Describir las características más sobresalientes de la producción escrita de aprendices adultos suecos de E/LE a través de identificar la frecuencia y el uso de combinaciones de palabras con *que*: *lo que*, *de que*, *algo que*, *dice que* en textos argumentativos recopilados en dos universidades suecas, durante los cursos lectivos 2008/09 y 2009/10.

(2) Comparar las características de la producción escrita de los aprendices adultos suecos de E/LE objeto del presente estudio con las reportadas en los dos corpus de

⁷ Véase en: http://cvc.cervantes.es/obref/marco/cap_03_01.htm. [Fecha de la consulta: 13-02-2010].

INTRODUCCIÓN

control cuyos textos han sido producidos por nativos hispanohablantes (subcorpus del CEDEL2) y nativos suecos (subcorpus del ARGUS).

(3) Realizar un análisis contrastivo del uso de las combinaciones de palabras con *que: lo que, de que, algo que, dice que* y sus equivalentes suecos. El análisis no trata aspectos de la traducción, sino de la interlengua para determinar si la función en contexto y frecuencia de uso puede considerarse idéntica o parecida en los textos producidos por aprendices suecos de español como lengua extranjera, comparado con el subcorpus del CEDEL2 (nativos hispanohablantes) y el subcorpus del ARGUS (nativos de lengua sueca).

(4) Compilar los datos de la encuesta aplicada a los participantes (suecos aprendices de E/LE) y analizar los resultados con el propósito de obtener información sociolingüística acerca de los parámetros: edad, sexo, tiempo de estudio de E/LE, lugar que han estudiado, frecuencia, otras lenguas estudiadas y tiempo que la han estudiado, así como su autovaloración sobre el dominio en las destrezas escritas.

(5) Analizar si hay correspondencia entre los descriptores del MCERL para el nivel A2-B1 y la competencia real de los aprendices del corpus SAELE.

El fundamento de estas preguntas de investigación se sustentan en la hipótesis de que los aprendices suecos de E/LE utilizan con una mayor frecuencia determinadas combinaciones de palabras con *que: lo que, de que, algo que, dice que* respecto a las correspondientes utilizadas por los nativos suecos e hispanohablantes de los corpus de control, sobreuso motivado por la influencia de la lengua materna (sueco) y a estrategias de acomodación de las cuales ellos se auxilian en su interlengua.

Las diferencias en formas y frecuencias de esas combinaciones en comparación con los hispanohablantes disminuyen proporcionalmente con la adquisición de mayores conocimientos y desarrollo de destrezas en la lengua meta.

1.2 Esquema metodológico

Animados por el propósito de estudiar de la forma más rigurosa posible el material disponible, responder a las preguntas de investigación y validar o no la hipótesis planteada hemos seguido el esquema metodológico que presentamos a continuación.

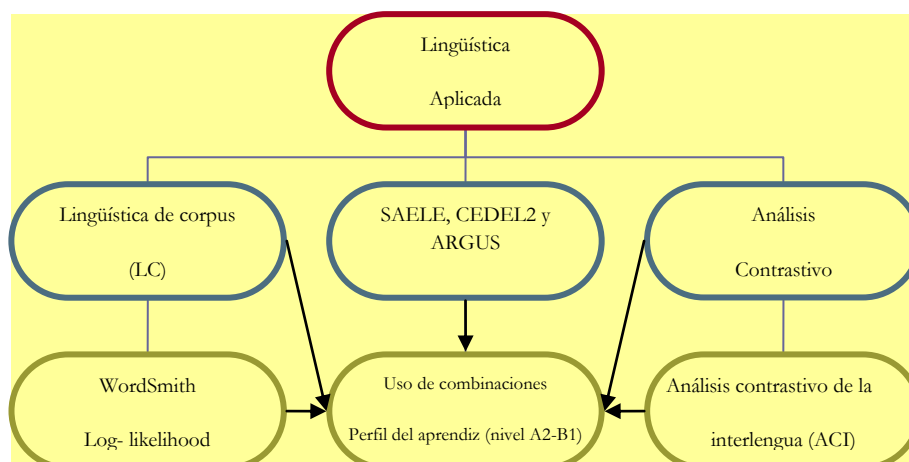


Diagrama 1. Esquema metodológico

Como se aprecia en el diagrama 1, la investigación se sustenta en la combinación de dos metodologías dentro de la lingüística aplicada, primeramente la lingüística de corpus (LC) que nos permite, con ayuda del procesador de texto WordSmith y el calculador Log-likelihood, realizar diferentes listas de co-ocurrencia de las combinaciones de palabras con *que*, entre ellas, la lista de las palabras del corpus con su frecuencia, la lista de colocaciones, validaciones comparativas de frecuencias, etc. En segundo lugar, la otra metodología es el análisis contrastivo (AC), y debajo de la misma colocamos el análisis contrastivo de la interlengua (en adelante, ACI), debido a que de acuerdo con nuestros objetivos pretendemos identificar eventuales correspondencias entre el uso de combinaciones de palabras con *que: lo que, de que, algo que, dice que* producidas por aprendices suecos de E/LE en comparación con las producidas por nativos suecos e hispanohablantes a través de un ACI.

Debajo de LA y en relación con ésta aparece el material que utilizamos: el corpus SAELE compilado en el presente estudio y dos corpus de control, uno de hispanohablantes (subcorpus del CEDEL2) y otro de nativos suecos (subcorpus del

ARGUS), ya que los participantes en nuestro estudio tienen el sueco como lengua materna y estudian español como lengua extranjera.

Paralelamente se ha aplicado una encuesta especialmente elaborada con la pretensión de revelar algunas características sociolingüísticas de los participantes permitiéndonos ganar en elementos sobre el perfil del aprendiz sueco de E/LE que ha formado parte de este estudio⁸, en el que además nos centramos en el análisis del uso y función de las combinaciones de palabras con *que: lo que, de que, algo que, dice que* en los textos que forman parte del corpus SAELE.

1.3 Disposición

La presente tesis está estructurada en ocho capítulos.

En el primer capítulo “Introducción” exponemos la motivación de nuestro estudio, presentando objetivos e hipótesis. Establecemos que el objetivo general es comparar el uso de algunas combinaciones de palabras de la producción escrita de aprendices adultos suecos de E/LE (nivel A2-B1 de acuerdo a la escala del MCERL⁹) con las producidas por nativos hispanohablantes y suecos. Con el fin de ilustrar más claramente qué hacemos y cómo lo hacemos utilizamos un diagrama con los pasos metodológicos seguidos en la confección del trabajo para terminar presentando la disposición de la tesis.

En el segundo y tercer capítulos se aborda el marco teórico. Específicamente en el segundo capítulo que denominamos “Panorama del desarrollo de la lingüística de corpus y los corpus de aprendices”, se discute la relevancia metodológica de la LC en el diseño y construcción de corpus computacionales, y específicamente del corpus SAELE, se ofrece un panorama histórico de los avances más novedosos destacando su vínculo con la pedagogía y los estudios de adquisición de lenguas para finalmente reseñar los aportes de investigadores escandinavos.

⁸ Cuando hablamos de ‘perfil del aprendiz’ nos referimos a las características más significativas que se han visto que poseen los participantes de este estudio a partir de los objetivos trazados.

⁹ Véase en: http://cvc.cervantes.es/obref/marco/cap_03_01.htm. [Fecha de consulta: 13-02-2010].

En el tercer capítulo titulado “Análisis contrastivo y enseñanza-aprendizaje de lenguas” se examina el AC en su relación con los estudios de adquisición de lenguas extranjeras, señalando nuestra posición respecto a las definiciones de análisis contrastivo, lengua materna, lengua extranjera e interlengua. Seguidamente, se exponen las características de la enseñanza-aprendizaje del español en Suecia, ejemplificando con el currículo de estudio del español como lengua extranjera de la universidad de Gotemburgo y la facultad de educación y comunicación de la universidad de Jönköping, destacando especialmente el papel que desempeñan las destrezas en la lengua, basada en los descriptores establecidos en el MCERL. A continuación presentamos las destrezas y competencias comunicativas según el MCERL con una aplicación de estos descriptores para el análisis que realizamos posteriormente del nivel de los aprendices participantes en el presente estudio.

En el capítulo cuarto “Combinaciones de palabras con *que*” se explica el uso y funcionalidad de combinaciones de palabras con *que: lo que, de que, algo que, dice que* y sus equivalentes suecos. Se analizan especialmente las similitudes y diferencias en uso y función de estas combinaciones de palabras en español y sus equivalentes en sueco y la variabilidad que poseen algunas de estas en ambas lenguas a través de ejemplos de los corpus de control CEDEL2 (subcorpus de nativos hispanohablantes) y ARGUS (subcorpus de nativos suecos). Esta descripción, unida a los resultados de estudios precedentes sobre combinaciones de palabras facilita el análisis del uso de las combinaciones de palabras con *que* en los textos recopilados en el corpus SAELE.

En el quinto capítulo “Material de trabajo” se presenta el corpus de aprendices suecos SAELE explicando sus criterios de selección y especificidades metodológicas inspiradas en las recomendaciones de Sinclair (2005) y Biber (1993) anteriormente aplicadas a corpus como el ICLE (Granger et al., 2002) y el CEDEL2 (Lozano, 2008) que nos sirven de modelos. Se presentan igualmente los corpus de control: el subcorpus de nativos hispanohablantes del CEDEL2 y el subcorpus de nativos suecos conformado del ARGUS. Asimismo se abordan los parámetros metodológicos de la encuesta y la información que ésta nos ofrece sobre las características sociolingüísticas de los aprendices para concluir refiriéndonos a los sujetos participantes en el estudio de las dos universidades suecas elegidas.

En el capítulo sexto “Resultados cuantitativos de los corpus” aparecen los resultados del corpus SAELE, los corpus de control CEDEL2 (subcorpus de nativos

INTRODUCCIÓN

hispanohablantes) y el ARGUS (subcorpus de nativos suecos). Los resultados se exponen con auxilio de diagramas, figuras y cuadros confeccionados con ayuda del procesador de texto WordSmith y el calculador estadístico Log-likelihood.

El capítulo séptimo contiene la “Discusión de los resultados” orientada a presentar un análisis cualitativo de los resultados obtenidos para dar respuesta a los objetivos formulados en el capítulo 1. El capítulo lo dividimos en los acápites 7.1-7.3.1, para abordar los cinco objetivos propuestos en la investigación. Se presenta finalmente el perfil del aprendiz sueco de E/LE participante en el estudio.

En el octavo capítulo aparecen las “Conclusiones” destacando las consecuencias lingüísticas y didácticas que se infieren del estudio y las futuras recomendaciones que se derivan de este.

Por último, en los anexos se recoge el plan de estudio de los diferentes cursos en la Universidad de Gotemburgo (*Grundkurs 1-30hp*, 2008/09 y 2009/10) y la facultad de educación y comunicación del instituto superior de Jönköping (*Spanska 1-30hp*, 2008/09 y 2009/10), la carta de presentación del estudio a los participantes, el certificado de consentimiento informado y la encuesta. Aparece además un índice de abreviaturas (antes del índice).

CAPÍTULO 2. PANORAMA DEL DESARROLLO DE LA LINGÜÍSTICA DE CORPUS Y LOS CORPUS DE APRENDICES

The most important achievement of corpus linguistics is undoubtedly that it has put the use of language at the centre of linguistics. In theoretical as well as practical approaches to language, computer corpora have placed linguistics on a firm empirical footing, emphasising the functional and communicative basis of language (Aijmer & Altenberg, 2004b: 1)

El presente capítulo se destina íntegramente al examen de la lingüística de corpus, una de las dos herramientas metodológicas principales del presente estudio (acápite 2.1), a la historia de su desarrollo y trascendencia en lengua hispana (acápite 2.1.2), a destacar su relevancia en los países escandinavos y en particular Suecia (2.1.3), y por último su aplicación a la pedagogía, resaltando la proliferación de corpus de aprendices en la última década (acápite 2.2).

2.1 La lingüística de corpus como herramienta metodológica

La realización de estudios que empleaban el análisis de la conversación y la etnografía de la comunicación como las principales metodologías dominó durante las décadas de los años 1960-90 (Lazaraton 2000: 179; Lazaraton 2003: 2). El desarrollo acelerado de la computación a partir de los años ochenta posibilitó la aparición de corpus computacionales para estudiar la producción oral y escrita de nativos y de aprendices de segundas y terceras lenguas, hecho que repercutirá en las investigaciones en lingüística aplicada (Axelsson, 2000; Axelsson & Westergren, 2002; Gaskell & Cobb, 2004; Granger, Dagneaux & Meunier, 2002; Lozano, 2008; Lüdeling, Kytöe & McEnery, 2008; Sánchez, 2001, 1995).

El concepto de lingüística de corpus comenzó a utilizarse con su significado actual, “as a linguistic methodology which is founded on the use of electronic collections of naturally occurring texts viz. Corpora” (Granger, 2005: 14), a partir de su aparición en *Corpus Linguistics I: Recent Developments in the Use of Computer Corpora*, en 1984 pero no ha sido considerada como un campo de estudio propiamente sino como herramienta para explorar aspectos del lenguaje que necesitan una explicación o

descripción más detallada, aplicable a la mayoría de las ramas de la lingüística (McEnery & Wilson, 2001: 2; Flowerdew, 2004).

El desarrollo de la ingeniería lingüística ha permitido la conformación de grandes corpus lingüísticos computacionales como el *Bank of English* (1997) y el *British National Corpus* (1995) permitiendo estudiar de forma sistemática el uso y funciones de la lengua y describir más acertadamente sus variaciones en diferentes registros (Altenberg, 2007; Altenberg & Granger 2001; Aijmer, 2009, 2002; Aijmer & Stenström, 2004a; Biber, 2006; Connor & Upton, 2004; Flowerdew, 2008; Gavioli, 2005; Nesselhauf, 2005; Verdaguer, 2004; Wray, 2002).

La LC dispone en la actualidad de registros sistematizados de textos que sobrepasan los 200.000.000 de palabras, y posibilitan analizar más objetivamente el rango de variabilidad de las expresiones lingüísticas y sus regularidades evidentes, al estudiar muestras estadísticamente representativas del lenguaje *per se*, en su estructura y en sus formas de usos cuantitativa, cualitativa y empíricamente, e incluso recientemente se encuentran en construcción algunos corpus de más de un billón de palabras, cuya compilación se hace utilizando la red de Internet (Mc Enery et al. 2006: 4-11; Parodi, 2010: 29).

El estudio de la fraseología basado en la LC ha catalizado la redefinición de aspectos de la teoría lingüística (Tognini-Bonelli, 2002), y propiciado un nuevo acercamiento a la investigación del lenguaje, no sólo como disciplina por derecho propio sino como ‘metodología’ general (McEnery et al., 2006: 6), criterio con el que coincidimos.

La LC establece criterios de garantía de la calidad metodológica en la conformación de un corpus, tales como tamaño y representatividad de la muestra, la anotación y la arquitectura e interfaz (Davies, 2008: 162; Sinclair, 2005), especialmente meritorios han sido en este sentido los diez principios de Sinclair¹⁰ que posibilitan chequear requisitos de calidad en el diseño y la elaboración de un corpus junto a las recomendaciones propuestas por Biber (1993). Estos son:

- El contenido del corpus (Sinclair, 2005:1).
- La representatividad del tipo de lengua que se vaya a estudiar y la elección de estudios longitudinales (durante un período de tiempo

¹⁰ Véase la propuesta original en: Sinclair, 2005:1-14. Estos principios son seguidos por Lozano en el diseño del CEDEL2 (Lozano, 2008).

prefijado) o transversales (en las cuales el aspecto cronológico no está estrictamente observado, pero las muestras se eligen en forma que asegura la similitud entre las muestras a comparar) (Sinclair, 2005: 2). Ello equivale a la obligatoriedad de observar pautas tales como el tamaño de la muestra, el origen o procedencia del material, tipo de texto, género y la longitud (Biber, 1993: 2-3).

- El ajuste entre el tamaño del corpus y las especificidades del fenómeno que se pretenda estudiar. Los corpus pequeños pueden ser perfectamente adecuados cuando el objetivo es investigar los elementos más comunes de la lengua. La representatividad está por tanto no sólo en relación con el tamaño de la muestra respecto al universo para el cual desean emitirse conclusiones y recomendaciones generalizadoras, sino también por los criterios de selección aplicados tales como para quién y qué objetivo persigue el mismo, criterios de inclusión/exclusión y forma de recolección de datos. El proceso deviene en realidad en una combinación de juicio y conveniencia en la elección de la muestra (Biber, 1993: 26).
- Elección de una estructura y plan diseñados que garanticen que los objetivos de la investigación queden validados en forma convincente tanto en la recolección de datos, como en su procesamiento e interpretación (Sinclair, 2005: 2, 5).
- Etiquetado. La información recopilada originalmente en los textos debe archivar de forma segura y aparte de los textos que se usan en el análisis para garantizar que eventuales necesidades de control o reelaboración dispongan de la información original recopilada directamente de la muestra de estudio. Lo que concretamente se etiqueta debe planificarse con anterioridad y responder a los objetivos específicos de cuáles características de la lengua se pretenden estudiar (Sinclair, 2005: 5).
- Autenticidad. Los documentos o transcripciones del habla serán completas e inalteradas (Sinclair, 2005: 7).

- Documentación de los hechos en el proceso de investigación en todas sus etapas. Esto permite que durante el análisis de los resultados, en caso necesario, se pueda hacer una comprobación accediendo a los datos originalmente copiados (Sinclair, 2005: 8).
- Equilibrio en la representatividad que garantice balance entre las muestras para evitar fuentes de errores en el análisis y la comparación. Principio especialmente decisivo en los llamados corpus monitor¹¹ (Sinclair, 2005: 9).
- El tema de redacción/conversación deberá ser representativo del registro del que se trate. Contenido y representatividad han de ser considerados (Sinclair, 2005: 10).
- La homogeneidad. Garantizar la correspondencia entre todos los componentes del corpus, eliminando textos atípicos u otros elementos que no respondan al diseño original del corpus (Sinclair, 2005: 14).

La forma en que estos principios han sido adaptados a los propósitos específicos de la conformación del corpus SAELE se explica en detalles en el Capítulo 5, y pretende alcanzar un diseño y rigurosidad de ejecución que pueda garantizar la validez de los resultados obtenidos siguiendo además el modelo del corpus CEDEL2 (Lozano, 2008).

¹¹*Monitor corpus is a corpus designed to track current changes in a language. A monitor corpus is added to annually, monthly or even daily, so it rapidly increases in size. However, the proportion of text types in the corpus remains constant, so that each year (or month or day) is directly comparable with every other (Hunston, 2006: 16).*

2.1.2 Análisis histórico del desarrollo de la lingüística de corpus

Hacia mediados de la década de 1950 se podía apreciar una tendencia en los lingüistas a trabajar recogiendo la mayor cantidad de datos posibles para sus estudios (McEnery, 2003: 452). En el primer estudio empírico aparecido en lengua inglesa (*Survey of English Usage*)¹² se recopiló a gran escala muestras de lengua para investigar la gramática. El estudio se convirtió en la gramática inglesa “*A Comprehensive Grammar of the English Language*”, la más importante durante muchas décadas (Quirk, Greenbaum, Leech & Svartvik, 1985), y sirvió de referencia a trabajos posteriores como el hoy conocido ICE (*Internacional Corpus of English*) (Teubert & Cermáková, 2007: 51).

Con la aparición en 1957 del libro *Syntactic Structures* de Chomsky, se criticó esta forma de compilar como poco objetiva (*skewed*) (McEnery & Wilson, 2001: 5; 1996: 4-8). La intuición como único criterio válido del estudio de la lengua y el papel central otorgado a la sintaxis en la primera etapa del generativismo, opuesto a los estructuralistas¹³ que desarrollaron sus trabajos fundamentalmente en el plano fonético-fonológico fueron los motivos declarados por este crítico (MacEnery & Wilson, 1996, 2001; McEnery, 2003; MacEnery, Xiao & Tono, 2006). Durante las dos décadas siguientes aparecieron solamente trabajos aislados en LC como el Corpus *Brown* (Francis y Kučera, 1964), que pretendía ser representativo del inglés americano escrito, su contrapartida inglesa el *LOB* (1978)¹⁴ y el *London-Lund Corpus* (1980)¹⁵ (Berber Sardinha, 2000: 330).

En este período en que la influencia del paradigma formalista primó en las investigaciones en la lengua inglesa, hubo progresos en el área de la lexicografía cuantitativa en lengua española y portuguesa con la aparición de un Diccionario de

¹² Fue compilado por Randolph Quirk y su equipo a partir de 1953 y convertido en corpus electrónico en 1989. Su parte hablada se computarizó antes y forma parte del London-Lund Corpus (Berber Sardinha, 2000: 324).

¹³ Los estructuralistas norteamericanos F. Boas; E. Sapir; L. Bloomfield y Ch. Fries se consideran iniciadores del empirismo y una rudimentaria LC (véase MacEnery & Wilson, 1996, 2001).

¹⁴ El Lancaster-Oslo-Bergen Corpus se desarrolló entre 1970-76, terminándose en 1978.

¹⁵ Jan Svartvik transcribió y digitalizó el Corpus Oral del *Survey of English Usage* (SEU) recogido por el lingüista británico Randolph Quirk a finales de los años 50.

frecuencia elaborado por A. Juilland¹⁶ & Chang-Rodríguez (1964) y el inicio del proyecto *Habla Culta* que abarcó los corpus de Juan Lópe Blanch (con material proveniente de 12 países de Hispanoamérica y España) y el *Linguagem Falada* de la lengua de Brasil (Davies, 2008: 149, 151; Fulk, 1993; Lópe, 1986).

La década de 1980 constituyó un vuelco positivo para el trabajo en LC, se conformaron grandes corpus lingüísticos basados en el desarrollo de las técnicas de computación. Labov (1969), Leech (1992) y Sinclair (1992) contribuyeron decisivamente mostrando las ventajas de estas herramientas respecto a la simple intuición, para observar de forma pragmática los datos, su frecuencia y permitir verificar los resultados (Davies, 2008; Tognini-Bonelli, 2002: 5; Sánchez, 1995: 17).

Las críticas sucesivas de Widdowson y Owen orientadas a minimizar el valor del corpus en la pedagogía ‘por no ser la lengua real sino solamente una muestra de producción del lenguaje’ fueron respondidas en el proyecto *Cobuild*¹⁷, fruto de la colaboración entre la Universidad de Birmingham y la Editorial Collins (más tarde HarperCollins) y el *Longman/Lancaster English Language Corpus* que demostraron el valor de la LC como herramienta para describir de forma más real la lengua con fines pedagógicos (Gavioli, 2005: 18; Moon, 2007: 179).

El proyecto *Cobuild*, comenzado en 1981, posibilitó –basado en un corpus- la publicación del *Collins Cobuild English Language Dictionary* (CCELD, 1986), acicate que revolucionó la lexicografía¹⁸(Moon, 2007: 160-163).

Durante este período, la ingeniería lingüística se desarrolló sucesivamente en España como respuesta a la necesidad de elevar la calidad en el trabajo académico y empresarial (Llisterri & Garrido, 1998; Llisterri, 2004). A principios de los 90 comenzaron en la Universidad Autónoma de Madrid trabajos que dieron a la luz el *Corpus Oral de Referencia del Español Contemporáneo*, el *Corpus de Referencia de la Lengua Española en Chile* (contentivo de un millón de palabras del español escrito en Chile) y el *Corpus de la lengua española en Argentina* (contentivo de un millón de palabras de la lengua escrita de Argentina). Apareció sucesivamente el *Habla Culta* (contentivo de

¹⁶ A. Juilland establece los conceptos de marco de muestra, representatividad y equilibrio.

¹⁷ Cobuild es el acrónimo de Collins Birmingham University International Language Database. The Bank of English Corpus. Véase en: www.cobuild.collins.co.uk

¹⁸ Véase el artículo de Rosamund Moon en la lista de referencias.

2.600.000 palabras extraídas de conversaciones de hablantes de Cuba, Puerto Rico, México, Costa Rica, Venezuela, Colombia, Perú, Chile, Bolivia, Argentina y España).

Estos cuatro corpus forman actualmente parte del CREA¹⁹, banco de datos que recopila derivaciones formales y variantes de la lengua literaria y no literaria en todos los países hispanohablantes con alrededor de 410 millones de palabras²⁰, creado a finales de los años noventa y que consta de una sección diacrónica que pretende ser una muestra representativa de esta lengua a lo largo de su historia (el *corpus diacrónico del español* o *CORDE*), y otra sincrónica dedicada a recoger el español de uso de los últimos años (el *corpus de referencia del español actual* o *CREA*) (Davies, 2008: 158).

El CORDE recopila 300 millones de palabras, abarca registros desde el inicio del idioma hasta el año 1975, y se divide en tres partes: Edad media, Siglo de oro y Época Contemporánea. Contiene muestras del español donde se habló y se habla, pero dada su perspectiva diacrónica contiene un 75 % del español de la península y un 25 % del resto de las variantes²¹. La versión actual del CREA consta de 154.279.050 formas, pertenecientes a textos procedentes de todos los países hispánicos producidos entre 1975-2004. En el 2008 se agregó un bloque básicamente con extractos de la prensa americana. Detalles acerca de número de textos, número de formas por países, temas y tipos de texto pueden obtenerse a través de la consulta de la nómina. Este corpus, considerado en Medellín, como el *corpus del Español del Siglo XXI*²², proporcionará materiales básicos a proyectos académicos en la investigación lexicográfica y gramatical del español en las veintiuna instituciones que forman la Asociación de Academias de la Lengua Española.

En el 2000 concluyó la compilación del *corpus CUMBRE*, patrocinado por la editorial SGEL, S.A., en España, contentivo de 20 millones de palabras extraídas de una

¹⁹ Acrónimo de Corpus de Referencia del Español Actual.

²⁰ El banco de datos es accesible a través de la página Web de la Real Academia de la Lengua Española: <http://www.rae.es/>

²¹ Información disponible en: http://Corpus.rae.es/ayuda_c.htm [24-10-2010].

²² Proyecto de la RAE, aprobado en el congreso de Academias (Medellín, 2007), según información que aparece en: <http://www.rae.es/rae/gestores/gespub000019>

variedad de textos del lenguaje oral y escrito, tanto del español peninsular como hispanoamericano²³ (Sánchez, 1995: 18).

El corpus *Archivo de Textos Hispánicos de la Universidad de Santiago*²⁴ contrasta 1.500.000 palabras procedentes del español contemporáneo con 1.063.969 procedentes del español medieval e incluye muestras de lenguaje oral y escrito producidos a través del desarrollo histórico de la lengua (Pérez, 2002).

Otros corpus de importancia son el *Corpus del español*, creado por Mark Davies²⁵ en la Universidad de Brigham Young y el proyecto PRESEEA, centrado en la creación de un corpus del español hablado, representativo del mundo hispánico en su variedad geográfica y social (Parodi, 2010: 159).²⁶

El vertiginoso desarrollo de la LC a partir de los noventa vinculada al progreso de la computación ha promovido una revolución cuantitativa y cualitativa en el área abriendo nuevas perspectivas al estudio del lenguaje (Tognini-Bonelli, 2001: 17), materializado por ejemplo en la creación de organizaciones como la asociación española AELINCO²⁷ (*Asociación de Lingüística de Corpus*), que ha celebrado varios congresos internacionales de lingüística de corpus (CILC 2009, 2010, 2011)²⁸ que permiten que investigadores de varios continentes presenten sus resultados en esta área del conocimiento (Cantos Gómez & Sánchez Pérez, 2009; Moskowich-Spiegel Fandiño et al. 2010), o a través de la prolífera actividad docente-investigativa en universidades como Lancaster, Birmingham y Lovaina.

A pesar de estos ejemplos, coincidimos con Davies en cuanto a que en lengua española son pocos los corpus online que cumplen al menos 2 de los cuatro principios de calidad mínima y sólo 5 corpus²⁹ siguen los requisitos que

²³ El *Gran diccionario de uso del español actual* (-GDUESA-, SGEL, 2001) en su versión gráfica y electrónica está basado en este corpus.

²⁴ Ya este Corpus forma parte del CORDE.

²⁵ Información obtenida en: <http://www.corpusdelespanol.org/>

²⁶ Información obtenida en: <http://www.linguas.net/Default.aspx?alias=www.linguas.net/portalpreseea>

²⁷ Véase la página de AELINCO en: <http://www.um.es/aelinco/>

²⁸ El IV congreso se acaba de realizar en marzo de 2012.

²⁹ El Corpus CREA, el CORDE, el VISL/CorpusEye, el Leeds Collection of Internet Corpora/Sketch Engine y el Corpus del Español.

consideramos imprescindibles: tamaño y representatividad de la muestra según el fenómeno sobre el que se pretenden emitir generalizaciones, aplicación de un sistema de anotación y arquitectura e interfaz (Davies, 2008:162).

En realidad opinamos que todos los principios propuestos (Sinclair, 2005) deben ser observados para que un corpus sea promovedor del conocimiento científico en el área a la que sus resultados pretenden orientarse.

En conclusión, podemos afirmar que los corpus posibilitan:

- Efectuar experimentos contrastables y replicables por otros investigadores desde perspectivas lingüísticas y didácticas precisadas (Baralo, 2010).
- Estudiar fenómenos de concordancia y variaciones de usos de formas gramaticales y lingüísticas en contextos específicos (Sánchez, 1995: 55).
- Producir listados de secuencias de palabras en contexto, ordenadas y clasificadas convenientemente siguiendo un formato homogéneo que facilitan identificar frecuencias de sus usos, concordancias, colocaciones y listas de palabras (Römer, 2010).
- Recopilar y estudiar características distintivas del discurso académico en comparación con otros géneros, mostrando por ejemplo que determinados sustantivos, las nominalizaciones, los sufijos predicativos y las conjunciones adverbiales son muy frecuentes en la prosa académica del inglés como lengua segunda (Gilkin, Granger & Paquot, 2007: 321).
- Optimizar el tiempo de recolección, organización, marcación y análisis de las categorías a explorar (Parodi, 2010: 167).

Estas ventajas de uso de la LC en el trabajo de exploración del lenguaje las hemos visto constatadas en la ejecución de la presente tesis, debido a que ha sido una herramienta para la mejor organización en la recolección del material, la presentación de los resultados y el análisis posterior de sus posibles significados lingüísticos.

2.1.3 Tradición e investigaciones en lingüística de corpus en Escandinavia y en Suecia en particular

El conocido y hoy computarizado *London-Lund Corpus*, iniciado en 1959, contó con la colaboración del lingüista sueco Svartvik junto a otros investigadores de la talla de Crystal, Greenbaum y Leech (Quirk et al. 1985). En la década de los 60 no se disponía aún de herramientas apropiadas para procesar digitalmente los textos y por ello no fue frecuente la confección de corpus con recopilación de cantidades de palabras que sobrepasaran el millón (Tognini-Bonelli, 2002: 17). No obstante se desarrollaron en Suecia los Corpus *Talsintax* con transcripciones de la lengua sueca hablada (sintaxis del habla), y *Skrivsintax* (sintaxis escrita) compuestos por textos en lengua sueca, tanto de alumnos de institutos como de escritores profesionales, que fueron entonces recopilados y etiquetados a mano.

Sture Allén es el pionero indiscutible de la lingüística de corpus en Suecia, quien durante su trabajo en la Universidad de Gotemburgo creó el *Press-65* contenido de 1.000.000 de palabras provenientes de la prensa escrita (matutinos o *morgontidningar*) y dio a la luz en 1975 el *Språkbanken*³⁰, que aún en la actualidad presta servicios a lingüistas de Suecia y otras partes del mundo para la realización de estudios empíricos de tipología o contrastivos sobre la lengua sueca y otras lenguas escandinavas.

El corpus *Nusvensk* (El sueco actual) sirvió de base a la conformación de un corpus de lengua escrita contemporánea, el corpus SUC(*Stockholm-Umeå Corpus*), que abarca más de 1.000.000 de palabras de diferentes géneros textuales³¹ (Lindberg, 1997).

En las décadas de los setenta y los ochenta la LC de tradición inglesa continuó creciendo en Escandinavia con la creación de corpus en Helsinki, Bergen, Oslo, Lund y Gotemburgo (Kopotev, 2003; McEnery & Wilson, 2001); muchos de los avances en los 70 se vincularon a Sture Allén, Knut Hofland, Stig Johansson y Jan Svartvik (Tognini-Bonelli, 2002: 16). Un producto de esa etapa en la colaboración de investigadores nórdicos y británicos fue la creación del LOB (Lancaster-Oslo-Bergen) (Teubert & Cermáková, 2007: 57).

³⁰*Språkbanken* (Centro de tecnología lingüística). Al momento de la realización del presente trabajo se ofrecía en la red el acceso al banco de datos Språkteknologi (Tecnología lingüística) que consta de doce entradas desde las cuales se puede acceder a diccionarios, compilaciones literarias y corpus de la lengua oral y escrita. <http://sprakteknologi.se/>

³¹*Göteborg Spoken Language Corpus* (Corpus de la lengua hablada de Gotemburgo) y el SUC (Stockholm-Umeå Corpus) antes mencionado. Véase en: <http://sprakteknologi.se/>

En la Universidad de Helsinki, se compiló el *The Helsinki Corpus of English Texts*³² por los investigadores Matti Rissanen, Merja Kytö y otros colaboradores en los últimos años del 80. Este corpus contiene textos de la lengua inglesa de diversas épocas y un total de 1.572.800 palabras. El *Oulu Corpus*, fue creado en el departamento de finlandés y saamian en la universidad de Oulu en el mismo período (Kopotev, 2003). Actualmente existen varios centros encargados de la recopilación de textos de lengua oral y escrita y con formación de corpus en Finlandia, entre los que puede mencionarse *The University of Helsinki Language Corpus Servers*, el *Finnish Center for Science*, el *Finnish Language Texts Bank*, el *Finland Language Research Center*, entre otros (Kopotev, 2003: 37-43).

La Universidad de Oslo ha sido también internacionalmente reconocida por el aporte de sus investigadores en los estudios de análisis contrastivo de lenguas con empleo de la LC, profundizando en aspectos de la semántica, la pragmática, la estilística y la traducción. Estos esfuerzos están patentes en el Proyecto SPRIK (*Språk i Kontrast/ Lenguas en contraste*)³³, destinado básicamente a la elaboración del corpus paralelo [Oslo Multilingual Corpus](#) (OMC), cuyo objetivo ha sido apoyar las investigaciones en traducción y aprendizaje de LE. Entre los lingüistas noruegos destaca Stig Johansson quien ha utilizado la LC en combinación con el análisis contrastivo para investigar aspectos sobre todo de la interacción entre el noruego y el inglés (Johansson, 2007; 2009).

The Bergen Corpus of London Teenage Language (COLT 1993) fue el primer corpus centrado en el estudio del lenguaje juvenil en lengua inglesa oral con una muestra de personas de 13 a 17 años de edad, en diferentes zonas de Londres. El corpus tiene un millón de palabras y forma parte del *British National Corpus*. En la actualidad existe el proyecto UNO (*Språkkontakt och ungdomsspråk i Norden/ Lenguas en contacto y lengua juvenil en los países nórdicos*) que estudia tanto las interrelaciones entre lenguas en

³² Información sobre el corpus: <http://khnt.hit.uib.no/icame/manuals/hc/INDEX>. HTM

³³ Todos los detalles del proyecto véanse en: <http://www.hf.uio.no/ilos/forskning/prosjekter/sprik/>

contacto como la lengua juvenil coloquial en Noruega, Islandia, Suecia, Dinamarca y Finlandia³⁴.

Al redactar esta tesis, existían en Suecia varios proyectos de LC en desarrollo en universidades como Uppsala, Lund, Estocolmo y Gotemburgo. El trabajo en la Universidad de Uppsala se concentra sobretodo en el departamento de inglés, básicamente con la compilación de corpus históricos sobre variación lingüística, corpus de aprendices para estudios de la IL (por ejemplo el ARGUS, que utilizamos como corpus de control en nuestra tesis fue realizado por una investigadora de esa universidad), análisis de errores en el aprendizaje y elaboración de manuales de enseñanza basados en estos resultados. Una novedosa línea de investigación compila en la actualidad un corpus de inglés utilizado como lenguaje en la comunicación computacional a través de internet.³⁵

Según McEnery & Wilson probablemente sea Bengt Altenberg (Universidad de Lund) el autor sueco más frecuentemente citado por investigadores extranjeros (McEnery & Wilson, 2001), quien a través del análisis contrastivo y la LC ha comparado la lengua sueca y la inglesa en estudios tales como *“Conclusive English ‘then’ and Swedish ‘då’. A corpus-based contrastive study”* (2010), *“The correspondence of resultive connectors in English and Swedish”* (2007) y *“Causative constructions in English and Swedish. A corpus-based contrastive study”* (2002). Karin Aijmer, de la universidad de Gotemburgo es también con gran frecuencia citada internacionalmente en el campo de LC y considerada pionera en la rama de la pragmática a través de sus estudios sobre conectividad y multifuncionalidad de marcadores en el discurso (Aijmer, 2009, 2006, 2004a, 2002).

En el Departamento de lengua y literaturas de la Universidad de Gotemburgo se realizan en la actualidad diversos proyectos basados en LC³⁶ que abarcan estudios en historia de la lengua, lexicología, traducción, pragmática y adquisición y enseñanza de lenguas. En esta última área está el de producción oral y escrita de aprendices suecos

³⁴ Página Web del proyecto *Språkkontakt og Ungdomsspråk i Norden* en: <http://www.uib.no/uno/>

³⁵ Véase en: <http://www.teknat.uu.se/forskning/uu/beskrivning.php?id=419&vetenskapsid=0&lang=sv>. Referido a la lengua española hay un proyecto en colaboración con las universidades de Alcalá de Henares y de Murcia (Corpus hispánico y americano en la red: textos antiguos). Véase en: <http://www.charta.es/>

³⁶ Existe una página informativa sobre investigadores y doctorandos que utilizan la LC como metodología en sus proyectos.

avanzados de inglés, basada en SWICLE (*The Swedish Component of the International Corpus of Learner English*) y LINDSEI-SW (*The Swedish Component of the Louvain International Database of Spoken English*) y estudios contrastivos y de traducción con la ayuda de *The English-Swedish Parallel Corpus* (ESPC). En el centro de tecnología lingüística de dicha universidad (*Språkbanken*) se recopilan diversos tipos de corpus de la lengua sueca producidos en diferentes épocas y géneros. Entre ellos un solo corpus de lengua española, el *Sol-Spanish online*³⁷ publicado en la red desde 1998 y resultado de una cooperación entre las universidades de Gotemburgo y Salamanca. Este corpus constituye un importante aporte a la aplicación de LC en estudios de la lengua española en la península escandinava.

2.2 Aplicación pedagógica de la LC y los corpus de aprendices

La observación del lenguaje en colecciones de textos de aprendices se inició en las décadas de 1960 y 1970, principalmente para el análisis de errores³⁸ (en adelante, AE) de aprendices de una L2 sin una sistematización del producto del análisis ni auxilio de la computación (Nesselhauf, 2005: 129). El término corpus de aprendices apareció por primera vez en el diccionario de Longman *Corpus Based Language Studie. An Advanced Resource Book*. (McEnery, Xiao & Tono, 2006: 248).

La aplicación de resultados de la LC a la pedagogía ha sido criticada por el hecho de que un corpus presenta la lengua fuera del contexto original (*Flowerden, 2009: 394*) y cuando se presentan las líneas de las concordancias, estas no muestran el discurso en su totalidad (Kaltenböck & Mehlmauer-Larcher, 2005: 71). Básicamente por estas razones se ha considerado no conveniente transferir datos directamente del corpus a los materiales pedagógicos a causa de la pérdida del contexto cultural en el cual los datos han sido recopilados, abogando por la necesidad de transformar las muestras

³⁷Spanish On-Line (SOL) tiene alrededor de 2.900.000 del español actual. Véase en:
<http://spraakbanken.gu.se/konk/rom2/>

³⁸ Modelo de investigación en L2, promovido por S. P. Corder, quien inspirado en la gramática generativa de Chomsky cuestiona el behaviorismo de Skinner y lleva a un replanteamiento tanto de la teoría del aprendizaje como del tratamiento de los errores desde una perspectiva más tolerante.

de acuerdo a la situación pedagógica (Braun, 2007; McCarthy 2001; Widdowson, 1991), consultar cuidadosamente la recopilaciones (Widdowson, 1998) y efectuar cierta mediación en que se adapten las muestras en dependencia de la situación pedagógica específica (Flowerdew, 2009: 404).

La aplicación de la LC en investigaciones de adquisición de lenguas extranjeras y lenguas segundas (L2) se impulsó en la década de 1980. Durante los noventa apareció el primer corpus de aprendices: el *International Corpus of Learner English* (ICLE)³⁹ (Granger, Dagneaux & Meunier, 2002: 14), seguido del *Louvain Corpus of Native English Essays* (LOCNESS) que pretendió ofrecer un corpus de control de inglés nativo con el cual contrastar la lengua meta y describir la interlengua. Un total de 104 corpus de aprendices creados en diferentes partes del mundo de variadas lenguas maternas aparecen recopilados en la página Web de la Universidad de Louvain⁴⁰.

Coincidimos con otros autores en que la creación de corpus de aprendices es sin dudas un campo novedoso en los estudios de la LA contemporánea (Alonso, 2007; Flowerdew, 2009; Granger, 2007; Torruella & Llisterri, 1999), que ha permitido un mejor análisis de errores a través de la observación de la lengua escrita, tipos de argumentaciones y formas de uso de los marcadores del discurso, revelando características de la IL (Barlow, 2005; Granger, 1998; Nesselhauf, 2004; Siepmann, 2005).

A pesar de que se han publicado relativamente pocos estudios empíricos que evalúen los resultados del uso de corpus en la enseñanza-aprendizaje (Kern, 2006: 193; Cobb, 1997: 301), y que los realizados tienen la limitación de basarse en muestras pequeñas de sujetos (Chambers, 2007: 5-6), se considera que la recopilación de corpus de aprendices constituye una base empírica importante para la creación de materiales didácticos adecuados al nivel del aprendiz que tomen en consideración los errores, las estrategias comunicativas de los estudiantes según sus niveles y detalles de la IL producida en la lengua meta (Torruella & Llisterri, 1999), constituyendo un recurso en las investigaciones de adquisición de lenguas segundas, específicamente en la adquisición de morfemas (Ellis, 2003).

³⁹ Está compuesto de textos producidos por aprendices de inglés como lengua segunda cuyas lenguas maternas son diferentes.

⁴⁰ Véase en: <http://www.uclouvain.be/en-cecl-icWorld.html>

Lo planteado anteriormente ha sido una de las principales motivaciones del presente estudio para tratar de aumentar nuestros conocimientos sobre las características de la producción escrita de los aprendices suecos de E/LE en camino a la lengua meta, para de esa forma recomendar aspectos específicos de la LE a los cuales debe prestarse mayor atención en el proceso de enseñanza-aprendizaje e incluso en un futuro trabajar en cómo estos deben ser presentados en los manuales o materiales didácticos que se utilicen.

Los aprendices han podido desarrollar habilidades en el uso en contexto de verbos de alta frecuencia gracias a la aplicación de ejercicios con corpus computacionales (Altenberg & Granger, 2001: 173-195) obtener auxilio complementario en el aprendizaje de la gramática (Estling & Lindquist, 2007: 329), desarrollar destrezas en la lengua escrita en LE y L2 (Gilmore, 2008: 3; Granger, 2007: 252) y trabajar con textos didácticos apropiados que describen aspectos tradicionalmente problemáticos en la adquisición de la lenguas (Gavioli, 2005: 25).

Estos estudios han posibilitado igualmente identificar aspectos de la IL de aprendices tales como la tendencia a utilizar más frecuentemente determinados vocablos y frases prefabricadas (Granger, 1998), el uso de combinaciones de palabras esencialmente típicas de la lengua oral, la tendencia a modificar adjetivos y subutilizar dispositivos de cobertura conllevando a la elaboración de una forma de escritura particular (Flowerdew, 2000: 151). A consecuencia de ello compartimos el criterio de que todo ajuste en la organización de la enseñanza basada en un mejor conocimiento de aquellas palabras y estructuras de la lengua meta que tienden a ser usados (en exceso, defecto o incorrectamente) por los aprendices es una aplicación práctica de notable valor en la enseñanza-aprendizaje de lenguas (Leech, 1997: 20).

Precisamente en lo anteriormente abordado reside una de las motivaciones en que se basan uno de los objetivos de nuestra investigación: determinar en qué proporción los aprendices suecos de E/LE utilizan las combinaciones de palabras con *que*: *lo que*, *de que*, *algo que*, *dice que* son utilizadas por los aprendices suecos de E/LE respecto a los participantes de los corpus de control para posteriormente analizar por qué lo hacen de esa manera (Cap.1).

El interés creciente por estudiar el proceso de adquisición del Español L2 ha estimulado la aparición de corpus como el *Spanish Learner Language Oral Corpus* (SPLLOC) que es un corpus oral L1 inglés-L2 español de estudiantes anglohablantes de la Universidad de Southampton, el cual utiliza tres tipos de tareas para los aprendices y compara los resultados con datos de un grupo de nativos hispanohablantes. Tiene como objetivo investigar el proceso de adquisición de propiedades morfosintácticas del español tales como orden de las palabras, pronombres clíticos, morfología verbal y formas interrogativas, atendiendo a las diferentes etapas de apropiación del E/LE (Baralo, 2010; Mitchell et al. 2008).

En la actualidad también se estaban desarrollando en España corpus de aprendices con el español como LM y el inglés como lengua meta tales como el *Written Corpus of Learner English* (WriCLE) de la Universidad Autónoma de Madrid y el *Santiago University Learner of English* (SULEC). Igualmente está en desarrollo el *Corpus Escrito del Español*⁴¹ (CEDEL2), que forma parte del proyecto *Word Order in Second Language Acquisition Corpora* (Chocano et al. 2007), compuesto por textos producidos por nativos de lengua inglesa que estudian E/LE. El interés de estos corpus se explica por la escasez de corpus de nativos de otras lenguas que aprenden el español como lengua extranjera (Davies, 2008: 181).

Otro corpus llamativo es el *Spanish Learner Corpus and Exercises* (SLCE) de la Universidad de Texas, que pretende familiarizar a los aprendices con el español como idioma meta en diferentes niveles de aprendizaje (Koike, 2007). Por otra parte, el *USP Multilingual Learner Corpus* (MLC) contiene producciones en español, entre otros idiomas, de alumnos de la Universidad de San Pablo, Brasil (Tagnin, 2002), mientras que el *European Science Foundation Second Language Databank* (ESFSLDB), es un corpus de conversaciones transcritas de aprendices para estudiar la IL en forma contrastiva a través del análisis de errores (AE) en la adquisición de múltiples idiomas como L2/LE, entre otras el español⁴².

Coincido con la opinión de Baralo, referida a que a medida que se amplía y se intensifica la investigación en la enseñanza y el aprendizaje de E/LE, surgen

⁴¹ Véase en: <http://www.uam.es/woslac/cedel2.htm>

⁴² Forma parte de la base de datos del Instituto Max Planck (<http://www.mpi.nl/world/tg/lapp/esf/esf.html>). En el proyecto han colaborado Francia, Alemania, Inglaterra, Holanda y Suecia.

preguntas de investigación que requieren ser abordadas a través de la descripción y el análisis de datos fiables como los que ofrecen los corpus de aprendices de E/LE. Ello abarca la identificación de tipos y frecuencias de errores, estandarización de criterios que permitan identificar el nivel de dominio de la lengua meta en que los aprendices se encuentran, la posible sistematización de tipos y frecuencias de errores respecto a etapas de desarrollo y características de la IL en el proceso de asimilación de la lengua meta dependiendo de la lengua materna del aprendiz y cómo esta última influye en el tipo de errores que el aprendiz comete y en la velocidad de adquisición de destrezas, así como identificar las estructuras sintácticas, fonéticas y gramáticas que deben trabajarse más intensiva o extensivamente (Baralo, 2010).

La aplicación de la LC como metodología unida al análisis contrastivo de la interlengua nos va a permitir registrar resultados cuantitativos y cualitativos fiables para poder validar la hipótesis de este estudio.

CAPÍTULO 3. ANÁLISIS CONTRASTIVO Y ENSEÑANZA- APRENDIZAJE DE LENGUAS

El análisis contrastivo es una de las metodologías que se utiliza en este estudio para analizar la interlengua de los aprendices suecos de E/LE y en este capítulo abordamos los aspectos principales de su contenido.

Para realizar el análisis de la interlengua de los aprendices suecos de E/LE es imprescindible examinar el proceso de enseñanza–aprendizaje del español como lengua extranjera en la universidad sueca, tratando de describir sus particularidades en la universidad de Gotemburgo y en el Instituto superior de Jönköping. El estudio se ha guiado por las indicaciones del MCERL referidas a niveles comunes de referencia, destrezas y competencias lingüísticas.

Además, puntualizamos conceptos básicos tales como lengua materna, lengua extranjera, lengua segunda, análisis contrastivo, interlengua y destrezas de lengua.

3.1 La lengua materna, la lengua extranjera y la interlengua del aprendiz de idioma

La lengua materna se considera base de la capacidad del individuo para adquirir una competencia lingüística y comunicativa determinada que le sirva para interactuar con sus semejantes, construir su mundo interior a través de sus percepciones, emociones, e inferencias, organizar su pensamiento, desarrollarse como ser humano (Baralo, 2004: 23; Cortés Moreno, 2000: 12) y la lengua que el individuo utiliza para comenzar a comunicarse (Gass & Selinker, 1993; Giacobbe, 1992), pero otros autores han considerado que diversos factores socio-contextuales influyen en que un individuo bi- o multilingüe emplee la L1 o la L2, en dependencia de situaciones comunicativas concretas (Postigo, 1998) y de funciones y usos de una u otra lengua en la comunidad del individuo multilingüe en cuestión, lo que determinará regularmente que una de ellas predomine en determinadas situaciones sobre la otra (Siguán & Mackey, 1986: 45).

El término lengua materna (LM) se ha utilizado indistintamente como lengua nativa, lengua de origen o lengua primera (en adelante, L1). Según el Diccionario de términos claves de E/LE del Centro Virtual Cervantes el término LM engloba tanto la LN como L1⁴³:

la primera lengua que aprende un ser humano en su infancia y que normalmente deviene su instrumento natural de pensamiento y comunicación. Con el mismo sentido también se emplea lengua nativa y, con menor frecuencia, lengua natal. Tal como se desprende de su apelativo, suele ser la lengua de la madre, aunque también puede ser la de cualquier otra persona: padre, abuelos, niñera, etc. Quienes comparten una lengua materna son considerados hablantes nativos de la lengua en cuestión, p. ej., hispanohablantes nativos

A los efectos del presente estudio consideramos que si bien la LM puede considerarse como lengua nativa y L1, la definición debe considerar la diversidad lingüística en el entorno del individuo plurilingüe, y por tanto la propia identificación de este con una de ellas como su LM en relación a su experiencia de uso en los diferentes contextos comunicativos. Para asegurar la homogeneidad de la muestra utilizada, hemos seguido un criterio de inclusión que tiene en cuenta que el aprendiz de E/LE de las universidades y niveles escogidos haya considerado el sueco como su LM, LN y L1.

El término L2 se ha asociado tradicionalmente con los estudios de adquisición de lenguas en su contexto natural de uso, y suele por tanto considerarse aquella que se aprende en comunidades donde subsisten al menos dos sistemas lingüísticos en contacto sea en situación de bilingüismo o de diglosia⁴⁴, o expresado en otros términos, la lengua no nativa adquirida siendo niño o joven en contexto natural (Baralo, 2004: 22-23) o, según Cook, la primera lengua aprendida como niño o joven, que no es la lengua materna (Cook, 2002: 1).

El término lengua extranjera (LE) se reserva para aquella adquirida generalmente como adulto en un contexto institucional (Baralo, 2004: 23), sobre la base de la LM (Porquier, 1991: 12). El E/LE es en el caso que nos ocupa una lengua estudiada

⁴³ Véase en: http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/lenguamaterna.htm

⁴⁴ Hace referencia a una situación de coexistencia de dos lenguas en una comunidad en que se produce un desequilibrio en la función social de uso de estas.

exclusivamente en forma de asignatura, no empleada comúnmente como medio de comunicación fuera del aula (Ellis, 2007) y es por tanto sinónimo de lengua meta y no precisamente L2, ya que constituye una parte del proceso enseñanza-aprendizaje en un país donde el español no es lengua materna (Gass & Selinker, 2008: 7), ni se emplea en forma natural por el aprendiz fuera de la clase⁴⁵. Estos últimos aspectos matizan el contexto situacional de nuestros sujetos según la teoría sociocultural debido a que en ésta se parte de que las personas actúan directa o indirectamente con otros en un marco interactivo práctico y un contexto cultural concreto (Säljö, 2000:104).

3.2 El análisis contrastivo de la interlengua

El rol que desempeña la LM en el proceso de adquisición y desarrollo de una L2 ha sido ampliamente abordado por lingüistas, psicólogos y pedagogos. En las décadas de 1940 y 1950 prevaleció el behaviorismo (modelo conductista) y con ello se concibió el aprendizaje de la lengua como un resultado de la imitación, la práctica, la retroalimentación (*feedback*) y los hábitos (Lado, 1957; Marín & Sánchez, 1991; Säljö, 2000).

Entre los años 1960 y 79 predominó la tendencia racionalista iniciada por Chomsky (1959), quien circunscribió el papel de la LM en la adquisición de L2 exclusivamente a las influencias determinadas por la transferencia de ciertos contrastes de la misma en relación con la lengua meta o la L2 (Gass, 1988: 384), otorgándole el papel decisivo a la predisposición del individuo desde edades tempranas para adquirir el lenguaje de la misma forma que cualesquiera otras funciones biológicas (Dulay & Burt, 1974), y concediendo máxima atención al estudio del proceso de pensamiento del individuo y sus capacidades o coeficiente (*ratio*). Durante este período florecieron los estudios sobre la interacción psicología-lingüística (Säljö, 2000: 55).

El análisis contrastivo (AC) devino en herramienta de análisis a partir de las teorías de Fries (1945), Haugen (1966), Lado (1957) y Weinreich (1953) orientado a predecir ciertas dificultades de los aprendices en la adquisición de la lengua meta, detectar diferencias y semejanzas entre la LN y la lengua meta en dos direcciones básicas:

⁴⁵ Aunque, como veremos en los resultados de la encuesta (capítulo 6), en nuestro corpus hay un número de aprendices que han estudiado en España o países latinoamericanos durante algún tiempo.

identificando las relaciones AC-materiales didácticos y técnicas pedagógicas en Norteamérica; o en Europa combinando el AC con la teoría lingüística para tratar de explicar mejor la naturaleza del lenguaje. Paulatinamente se orientó a identificar los errores de los aprendices como desviación en la lengua meta debida a interferencias de la LM, surgiendo así el denominado análisis de errores (en adelante, AE), modelo propuesto por Corder (1967) que reconoció el error como algo inevitable y positivo que evidenciaría que el proceso de aprendizaje había comenzado (Gass, 1988: 385).

Sucesivamente se estableció la necesidad del estudio empírico de la producción global del estudiante de una L2 con el objetivo de caracterizar el sistema lingüístico que éste emplea en el llamado análisis de la IL (Selinker, 1969, 1972) (en Santos Gargallo, 1993: 157). Según lo plantea George (1972) a principios de la década del sesenta el error era asociado con la corrección y a finales de la década con el proceso de aprendizaje.

A partir de los estudios de Selinker se generalizó el uso del término IL, que en la actualidad se considera alternativamente bien como un sistema estructurado construido por el aprendiz de una LE en estadios determinados del proceso de aprendizaje, o como una serie de sistemas entrelazados que conforman el denominado *continuum*⁴⁶ interlingüístico (Baralo, 2004: 35-39; Cook, 2002: 10). La interlengua debe considerarse como un sistema dinámico (en continuo desarrollo), en el proceso de adquisición y desarrollo de secuencias auxiliares en la comunicación en una lengua meta influida continua, pero no obligatoriamente de modo negativo, por la LM del aprendiz (Pastor Cesteros, 2005: 371)⁴⁷.

Durante la década de los 60 apareció también el término *transfer* en trabajos influidos por la tendencia Behaviorista o Conductista (Ringbom, 2007:30). A partir de la *Conference on Language Transfer* (Michigan, 1981), el término se redefinió como *cross-linguistic influence* (CLI) (Kellerman & Smith, 1986), y los estudios de adquisición de lenguas se ampliaron a nuevas temáticas como el contexto sociolingüístico del aprendiz, los sentimientos del aprendiz acerca de su propio aprendizaje (Odling, 1989), la relación entre la gramática universal y la L1 en la adquisición de una L2

⁴⁶ El *continuum* interlingüístico se observa en el proceso de separación, interconexión e integración entre la L1 y L2 o LE en el proceso de adquisición de una lengua.

⁴⁷ La autora resume las conclusiones ofrecidas por Larsen-Freeman & Long (1991).

(Epstein, Flynn, & Martohardjono, 1996; Hawkin, 2001) y la interferencia léxica (Helms-Park, 2001).

Es práctica habitual en publicaciones contemporáneas de estudios contrastivos de lenguas el utilizar el término *cross-linguistic influence* indistintamente como sinónimo de *language transfer* o *linguistic interference* para abordar variados aspectos sobre el papel de la LM, LN en la denominada mezcla de lenguajes o *language mixing* (Gass, 1988: 385; Odling, 2003: 436; Ringbom, 2007: 30), sin embargo en nuestra opinión la IL debe verse como un proceso de interacción permanente entre la lengua meta, la LM y otra u otras LE paralelamente adquiridas por el aprendiz (L2, L3, etc.). Las influencias son en todas las direcciones aunque a los efectos de los estudios de adquisición de lenguas se haya prestado ante todo atención a la influencia de las lenguas previas en la lengua meta. No existen en realidad ‘influencias negativas’ o ‘influencias positivas’, pero puede ser importante en el proceso de apropiación de la lengua meta apoyar el proceso de concientización en el aprendiz de aquellos conocimientos adquiridos y habilidades lingüísticas desarrolladas en uno o varios idiomas previamente estudiados que influyen en el proceso de apropiación de la lengua meta.

Por estas razones utilizamos el término influencia interlingüística a los efectos del presente estudio exclusivamente para describir la interacción entre similitudes y diferencias entre todas las lenguas adquiridas en etapa anterior a la lengua meta y el resultado final de dichas interacciones en el proceso de apropiación de esta última (Odling, 1989: 27). Ello equivale a considerar el AC como herramienta que compara de forma sistemática dos o más lenguas, para tratar de describir y detectar eventuales similitudes y diferencias entre estas (Johansson, 2007: 23). Esta forma de comprender la interlengua y aplicar el AC como herramienta es la base del llamado análisis contrastivo de la interlengua (en adelante, ACI).

Tradicionalmente se aplica el AC para el estudio de la IL contrastando un determinado fenómeno en dos lenguas. De esta manera se acostumbra a yuxtaponer el fenómeno lingüístico en cuestión para comprobar la equivalencia de estructuras y compararlas en cuanto a reglas sintácticas tratando de identificar construcciones idénticas (Krzyszowski, 1990; Söhrman, 2007:16). En la actualidad el AC, también denominado lingüística contrastiva, ha tenido numerosos avances, a partir del impacto de la computación, y más específicamente debido a que la LC la ha revolucionado (Aijmer, 2006: 5). Esto se observa en áreas tales como el procesamiento del lenguaje natural, la traducción y la lexicografía computarizadas,

donde la información obtenida a partir de los corpus devienen en una realidad científicamente más fiable que la derivada de la simple introspección y el razonamiento del investigador (Altenberg & Granger, 2002:5)

Es por estas razones que en este estudio adaptamos el modelo de ACI de Granger (Granger, 1996: 43, 2003, 2007; Gilquin, 2005, 2008, Paquot, 2008), para analizar la IL de los aprendices suecos de E/LE valiéndonos de la comparación entre el sueco (LM de los aprendices) y el español (LE y la lengua meta), combinado con la metodología de lingüística de corpus (Altenberg, 2007; Butler, 2008; Johansson, 2007; Siepman, 2005; Tizón Couto, 2009). En las comparaciones de la interlengua se juxtapone el fenómeno lingüístico analizado a construcciones equivalentes en la LM y la LE, tratando de identificar no sólo construcciones idénticas, erróneas y diferentes, sino incluso analizar los motivos de ello, como se aprecia en la figura 2. Concretamente en nuestra propuesta, comparamos el uso de combinaciones de palabras con *que*: *lo que*, *de que*, *algo que*, *dice que* en las lenguas sueca y española, para después medir estos resultados cuantitativa y cualitativamente con los de los aprendices suecos de E/LE (IL de los aprendices).

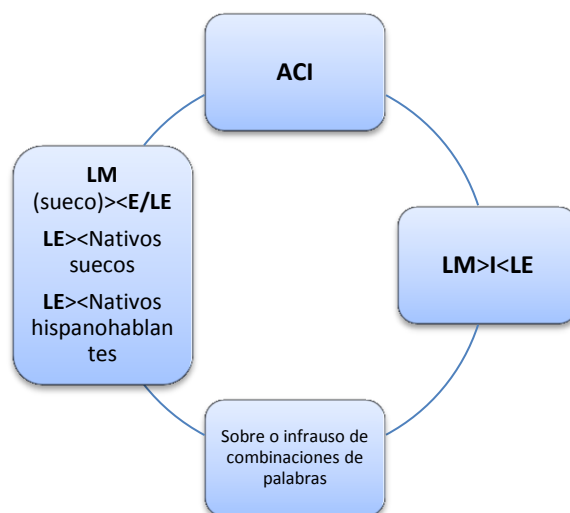


Figura 2. Propuesta metodológica del ACI

Los estudios de relación interlingüística entre la lengua meta y la LM han prestado atención a aspectos tales como la formación de secuencias naturales en el proceso de

adquisición, la relación de ambas lenguas entre sí y con el lenguaje universal, relaciones con el *markedness*⁴⁸ o *marcado*, el vínculo biunívoco entre L1 y L2, tendencias a evitar producciones difíciles determinadas (Laufer, 2000), sobre y/o infraproducción de determinadas estructuras (Odling, 1989: 80), el papel de la influencia interlingüística en el proceso de adquisición y aprendizaje de una tercera lengua (Bardel, 2000; Ringbom, 2007), entre otros. Un ejemplo de ello es cómo hablantes nativos de alemán cuya L2 es el inglés, suelen producir en francés L3 combinaciones más o menos fijas de vocablos como: *les gens sont impliqués*⁴⁹ (Dewaele, 1998); mientras que hablantes nativos de inglés con el alemán como L2 suelen producir en sueco L3 la palabra *föreslagger*⁵⁰ (Herwing, 2001).

La influencia de la LM sobre la L2 u otra LE puede determinar usos en defecto o exceso de determinadas estructuras por los aprendices (Altenberg, 2007; Jarvis, 1998, 2000; Tizón Couto, 2009), resultando en ocasiones la emisión de locuciones que – determinadas por una cierta maestría del aprendiz en el dominio de la gramática y el vocabulario de una determinada L2 pueden en apariencia ser fluidas para el aprendiz en dicha L2, pero que en realidad carecen de competencia pragmática, es decir constituyen locuciones que en el plano social y cultural no resultan las más apropiadas. En la comunicación intercultural, este aspecto se conoce como un punto distintivo de lo ‘no nativo’ (Beebe, Takahashi & Uliz-Weltz, 1990). Un ejemplo de ello, detectado en el marco sociocultural del presente estudio lo constituye el hecho que los nativos suecos aprendices de E/LE presentan dificultades en el uso de los verbos *ser-estar* (*to be* en inglés; *vara* en sueco), y aún en niveles avanzados de comunicación pueden elegir la variante inapropiada a determinado contexto en forma bastante ‘fluida’.

El AC ha demostrado ser una valiosa herramienta para localizar las zonas de interferencia interlingüística en las cuales la LM deviene en una de las principales fuentes de dificultad en la adquisición de L2. Sin embargo muchos errores derivan de las estrategias empleadas por los estudiantes en la adquisición del lenguaje y de la

⁴⁸ Según Ellis el término se refiere: “to the idea that some linguistic structures are ‘special’ or ‘less natural’ or ‘less basic’ than others. For example, the use of ‘break’ in ‘she broke my heart’ can be considered marked in relation to the use of ‘break’ in ‘she broke a cup’. Markedness can help to account for patterns of attested L2 acquisition” (Ellis, 2003:713).

⁴⁹ En inglés: *involved*; en francés: *impliqués*.

⁵⁰ En alemán: *vorschlagen*; en sueco: *föreslår*.

interferencia mutua de los artículos contenidos en el idioma de destino, y claro que estos no pueden ser estudiados con aplicación del AC (Richards, 1971).

En nuestro estudio⁵¹ empleamos la metodología de análisis de errores (en adelante AE) exclusivamente para tratar de describir y analizar algunos errores que los aprendices suecos de E/LE cometen al utilizar en sus textos combinaciones de palabras con *que* en el contexto fraseológico.

3.3 El proceso de enseñanza-aprendizaje de E/LE en la universidad de Gotemburgo y en el Instituto superior de Jönköping en Suecia

En el 2001 se estudiaban 35 lenguas modernas diferentes en las universidades e institutos superiores suecos (equivalente al 5% del total de asignaturas universitarias). El número de estudiantes universitarios matriculados en alguno de estos cursos había disminuido notablemente y la mitad de ellos estudiaban inglés. El descenso más dramático se registró en alemán y francés, con un crecimiento incipiente en español e italiano (HSV, 2001:19)⁵². En 2004 el español se había extendido en cuanto a lugares de enseñanza y número de estudiantes, convirtiéndose en la lengua romance más ampliamente representada; a pesar de ello disponía aún de menos recursos financieros asignados que el francés y el alemán⁵³. En 2010, se constataba la expansión sostenida del E/LE que ahora se impartía en 14 universidades e Institutos superiores a nivel de pregrado⁵⁴ y en 5 de ellas incluso a nivel de postgrado e investigación. A pesar de ello los recursos financieros seguían concentrados en el francés y el alemán (HSV, 2010:6)⁵⁵.

⁵¹ Véase los objetivos capítulo 1.

⁵² Högskoleverket (HSV), equivalente a Dirección nacional de educación superior sueca. Véase en: <http://www.hsv.se/download/18.539a949110f3d5914ec800090615/isbn91-88874-71-0.pdf-2009-12-05> [05-12-2009]

⁵³ Véase en: HSV, 2004. <http://www.hsv.se/download/18.539a949110f3d5914ec800092239/0503R-2.pdf#search=%27spanska%27-> [05-12-2009]

⁵⁴ Me refiero a las nueve instituciones que participan en este informe de la Agencia sueca de educación superior.

⁵⁵ Véase en: Rapport 2010:6 R <http://www.hsv.se/download/18.1dfeedbc1284e5ecc3d8000350/1006R-uppfoljning-utvarderingar-2004.pdf> [06-12-2010].

En este período de estudio se ofrecían cursos de E/LE en 15 universidades e Institutos superiores⁵⁶. A saber: Göteborgs universitet; Höskolan Dalarna; Höskolan i Halmstad; Höskolan i Jönköping; Karlstads universitet; Kungliga Tekniska högskolan; Linköpings universitet; Linnéuniversitetet; Luleå tekniska universitet; Lunds universitet; Malmö högskola; Mälardalens högskola; Stockholms universitet; Umeå universitet; Uppsala universitet.

El requisito de nivel de estudios precedentes del E/LE para ingresar en dichos cursos debe corresponder aproximadamente con el denominado nivel A2 del MCERL. El estudiante que es un usuario básico de la lengua debe ser capaz de:

Nivel A2 (Instituto Cervantes, 2002)

1. Comprender frases y expresiones de uso frecuente relacionadas con áreas de experiencia que le son especialmente relevantes (información básica sobre sí mismo y su familia, compras, lugares de interés, ocupaciones, etc.)
2. Comunicarse a la hora de llevar a cabo tareas simples y cotidianas que no requieran más que intercambios sencillos y directos de información sobre cuestiones que le son conocidas o habituales.
3. Describir en términos sencillos aspectos de su pasado y su entorno así como cuestiones relacionadas con sus necesidades inmediatas.

3.3.1 El contexto situacional del estudio

El pensamiento, la comunicación y las acciones físicas se sitúan en contextos relacionados entre sí y la perspectiva socio-cultural ve el vínculo entre el contexto y la acción individual en el proceso de aprendizaje (Säljö, 2000: 130). Partiendo de esta premisa es que presentamos a continuación detalles sobre el contexto educacional de

⁵⁶ En Suecia hay 14 universidades estatales y 21 Institutos superiores. Además hay 3 Institutos superiores que se subordinan a un privado y tienen autoridad para ofrecer programas de doctorados, siendo estos: Chalmers tekniska högskola, Handelshögskolan i Stockholm y Höskolan i Jönköping. Otros 9 Institutos superiores tienen autorización para impartir cursos de nivel básico o determinadas carreras de nivel medio superior. Véase: <http://www.hsv.se/densvenskahogskolan/universitetochhogskolor/adresser.4.539a949110f3d5914ec800062390.html> [06-12-2010].

los aprendices suecos de E/LE, participantes en este estudio para tener elementos sobre dónde y cómo han participado en el proceso de enseñanza-aprendizaje de E/LE.

INSTITUCIÓN Y ASIGNATURA: Durante los cursos lectivos que realizamos la recopilación del material, aunque el E/LE era una asignatura del plan de estudios en ambos centros de educación superior (Gotemburgo y Jönköping), no ocupaba el mismo lugar en dichas instituciones. En la universidad de Gotemburgo estaba incluido en el currículo desde nivel básico (1-60 créditos)⁵⁷ hasta nivel de doctorado, formaba parte de una sección del departamento de lenguas y literaturas de la facultad de humanidades y ofrecía cursos libres incluidos en el programa de profesores, cursos de máster, y cursos con fines específicos (complementarios en la formación de economistas y de traductores). Los diferentes cursos estaban organizados en dos alternativas temporales (a tiempo completo o medio tiempo) y espaciales (presenciales o a distancia).

En el instituto superior de Jönköping, por su parte, el E/LE se ofrecía exclusivamente como asignatura del currículo de estudios de cursos libres de la Facultad de educación y comunicación, en el nivel básico (1-60 créditos). La asignatura formaba parte de la sección de lenguas. Los cursos se ofrecían sólo a medio tiempo y con carácter presencial.

PROGRAMA DE ESTUDIOS: Durante los cursos lectivos 2008/09 y 2009/10 fueron impartidos 10 cursos de E/LE en el nivel básico (primer año) en la universidad de Gotemburgo y solamente cuatro en la Facultad de educación y comunicación del Instituto Superior de Jönköping. El estudio que realizamos se circunscribe a los textos producidos por los sujetos de la muestra en los cursos de destrezas que integraban respectivamente el curso *Spanska 1-30 hp*⁵⁸ (Jönköping) y el

⁵⁷ Llamado en sueco *Grundkurs I*. Para aquellos estudiantes que cumplan los requisitos de entrada al E/LE, pero que no se sienten con los conocimientos necesarios existe la posibilidad de realizar el curso introductorio denominado Español práctico o en sueco *Praktisk spanska*.

⁵⁸ Véase el plan de estudio correspondiente a los cursos lectivos 2008-09 y 2009-10 en el anexo 2. El currículo de estudios correspondiente al curso de 1-30 créditos está formado por cuatro cursos de 7,5 créditos cada uno, con un encuentro semanal de 3 horas durante 8 semanas. El primer curso era de Destreza oral y fonética, el segundo de Destreza escrita y gramática, el tercero de Cultura y texto no literario y el cuarto de Literatura hispanoamericana contemporánea. Los alumnos participantes en el presente estudio, cursaban el segundo curso, o sea el referido a Destreza escrita y gramática. El curso tenía un valor de 7,5 créditos. Entre los objetivos de este curso, que es el que nos atañe, estaba el desarrollar la destreza escrita a través de la lectura, la interpretación y la escritura en actividades

*Grundkurs I 1-30 hp*⁵⁹ (Gotemburgo) y debido a esto explicaremos solamente el plan de estudio correspondiente a los mismos.

PLANIFICACIÓN DE LA ASIGNATURA: Los contenidos de estos cursos, independientemente de que se desarrollaron en uno u otro centro de educación superior tuvieron un carácter similar en cuanto a su enfoque comunicativo centrado en las necesidades de los alumnos, observar una estrategia didáctica de trabajo común, emplear materiales, temas, actividades y formas de evaluación parecidas, con independencia a que los cursos no eran iguales.

Entre los objetivos del curso de destrezas impartidos en ambas universidades estuvo el desarrollar la destreza tanto oral como escrita. El enfoque utilizado combinó tanto el trabajo individual, en parejas como en grupos, tanto en el aula como fuera de ella. La mayoría de las tareas de comprensión y expresión escrita fueron realizadas en casa, en temas seleccionados previamente por los profesores⁶⁰: ‘Mi peli favorita’, ‘Vacaciones inolvidables’, ‘El tiempo libre’, ‘El deporte que me gusta practicar’, ‘Recuerdos de Navidad’ o bien argumentaciones sobre artículos periodísticos de actualidad. Redacciones de otros tipos de texto como descripciones y resúmenes, formaron también parte de las tareas asignadas a los sujetos pero no han sido considerados a los efectos del presente trabajo.

Los textos escritos y recopilados para la conformación del corpus SAELE, fueron producidos por algunos participantes de este curso, quienes los enviaron de forma voluntaria. Para la comunicación profesor-alumno, alumno-alumno y entrega de trabajos se empleó básicamente el contacto en la clase y el correo electrónico durante el curso lectivo 2008-2009 en ambos centros de educación superior; además fueron utilizadas las plataformas electrónicas GUL (Gotemburgo) y Ping Pong (Jönköping) durante el curso 2009-10 para el correo electrónico, comunicación Pim y envío de trabajos.

comunicativas de comprensión textual y la redacción de textos cortos donde se aplicasen los conocimientos de gramática adquiridos.

⁵⁹ Véase el plan de estudio correspondiente a los cursos lectivos 2008/09 y 2009/10 en el anexo 1. Este curso se componía, al momento de realizar nuestro trabajo, de varios pequeños cursos: gramática general; fonología; gramática española; destreza oral (en un solo curso), destreza escrita y texto, cultura y sociedad. El curso de destrezas tenía un valor de 5 créditos. y su objetivo principal era que el alumno desarrollase las destrezas necesarias en la lengua meta y de esa manera alcanzar el nivel B1 del MCERL. El curso tenía un encuentro semanal de 2 horas durante el semestre de estudios.

⁶⁰ Los temas fueron seleccionados, después de previa discusión, por el profesor Eduardo Jiménez Tornatore y la autora de esta tesis.

Los materiales de trabajo utilizados fueron: ejercicios seleccionados del libro de texto de *Destrezas integradas* de Nuria Sánchez Quintana & David Clark (1995), diversos textos filmicos, artículos periodísticos, canciones y textos literarios cortos. En cuanto a las actividades se pretendió hacerlas lo más variadas posibles, por ejemplo debate de artículos periodísticos, noticias y otros textos, ejercicios de comprensión auditiva, dramatizaciones en parejas o grupos. Los temas de las clases fueron: la música, la educación, España en cifras, el cine, la inmigración, el hogar, vacaciones, el empleo y los jóvenes, la literatura y el tiempo libre y el deporte. Como parte de la evaluación del curso los estudiantes debieron realizar una exposición final de tema libre sobre España o Latinoamérica. La presentación podía ser individual o en parejas y tenía una duración de cinco minutos por estudiante. Una profesora visitante participó en determinadas actividades en ambos centros de educación superior⁶¹, durante el otoño de 2009.

3.4 Las destrezas según el MCERL

Las denominadas destrezas idiomáticas, habilidades, capacidades comunicativas o macrohabilidades (Cassany, Sanz & Luna, 2007: 88), suelen tradicionalmente dividirse en dependencia al modo de uso de la lengua (oral o escrita) y a su papel en el proceso de comunicación (receptivo o productivo), clasificándose en: escuchar y leer; respectivamente hablar y escribir (Alcón Soler, 2002; Cortés Moreno, 2000). La lengua oral se considera más coloquial, subjetiva, redundante y abierta, de sintaxis más simple, emplea frases inacabadas, elipsis, repeticiones, un léxico más general y pobre, onomatopeyas y frases hechas. La lengua escrita suele ser estándar, más objetiva, precisa, cerrada, contiene un léxico específico y evita las repeticiones y la expresividad de los recursos lingüísticos populares (Cassany, Sanz & Luna, 2007: 90).

Como habíamos referido nos proponemos analizar cómo aprendices suecos que estudian E/LE en dos centros de enseñanza superior en Suecia utilizan las combinaciones de palabras con *que* en la producción de textos escritos.

Esclarecer el posible interés de ello presupone comprender el papel que se le otorga a las destrezas en el Marco Común Europeo de Referencias para las Lenguas,

⁶¹ Dr. Marisol Martínez. Universidad de Cienfuegos, Cuba.

especialmente a la relación destrezas escritas/competencia lingüística, al concebir más íntegramente los procesos de comprensión, expresión, interacción y mediación en la interacción oral y escrita en el proceso docente-educativo.

Para el MCERL son las competencias generales (individuales) y las competencias comunicativas de la lengua las que posibilitan al individuo actuar utilizando medios lingüísticos en un contexto y ámbito determinados, y adoptar una estrategia que le posibilita acometer una determinada tarea (Instituto Cervantes, 2002: 9).

Las relaciones entre destrezas y competencias en la lengua se clasifican en: generales (referidas a cualidades cognitivas, afectivas y conductuales del individuo que influyen en el ‘saber hacer’ y las comunicativas referidas a las competencias lingüísticas, sociolingüísticas y pragmáticas que influyen en el ‘saber crear’ (Instituto Cervantes, 2002: 99-27; González Piñeiro, Guillén Díaz & Vez, 2010: 49).

De ahí que veamos la relación entre ‘saber hacer’ de la destreza escrita y las competencias lingüísticas de tipo ‘saber crear’ en textos producidos por nativos suecos aprendices de E/LE, según lo planteado en el MCERL y siguiendo las propuestas de González Piñeiro (MCERL, cap.4-5; González Piñeiro et al., 2010: 115), señalando que se parte de un enfoque comunicativo en el proceso de enseñanza-aprendizaje y que en las actividades se han combinado todas las destrezas de la lengua⁶².

Reflexionando acerca de las descripciones precisadas en el MCERL es preciso afirmar que los nativos suecos aprendices de E/LE, sujetos del presente estudio se encontraban en un nivel aproximado de A2-B1 según la escala de niveles comunes de referencia del MCERL.

La competencia lingüística formal de nuestros aprendices se corresponde, de manera general con las destrezas que se especifican a continuación:

A2

⁶² Con independencia a que nuestro objetivo principal está centrado en la escrita.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

- Tiene un repertorio de elementos lingüísticos básicos que le permite abordar situaciones cotidianas de contenido predecible, aunque generalmente tiene que adaptar el mensaje y buscar palabras.
- Produce expresiones breves y habituales con el fin de satisfacer necesidades sencillas y concretas: datos personales, acciones habituales, carencias y necesidades, demandas de información.
- Utiliza estructuras sintácticas básicas y se comunica mediante frases, grupos de palabras y fórmulas memorizadas al referirse a sí mismo y a otras personas, a lo que hace, a los lugares, a las posesiones, etc.
- Tiene un repertorio limitado de frases cortas memorizadas que incluye situaciones predecibles de supervivencia; suelen incurrir en malentendidos e interrupciones si se trata de una situación de comunicación poco frecuente.

B1

- Dispone de suficientes elementos lingüísticos como para describir situaciones impredecibles, para explicar los puntos principales de una idea o un problema con razonable precisión y para expresar pensamientos sobre temas abstractos o culturales tales como la música y las películas.
- Dispone de suficientes elementos lingüísticos como para desenvolverse y suficiente vocabulario para expresarse con algunas dudas y circunloquios sobre temas como la familia, aficiones e intereses, trabajo, viajes y hechos de actualidad, pero las limitaciones léxicas provocan repeticiones e incluso, a veces, dificultades en la formulación.

En las actividades de expresión escrita, el usuario de la lengua en el nivel B1 debe producir un texto escrito que, según el MCERL, le permita formular redacciones cortas y sencillas sobre temas de interés; resumir, comunicar y ofrecer su opinión con cierta seguridad sobre hechos concretos relativos a asuntos cotidianos, habituales o no, propios de su especialidad; escribir informes muy breves en formato convencional con información sobre hechos habituales y los motivos de ciertas acciones (MCERL, cap. 5).

3.4.1 La destreza escrita y las competencias comunicativas

Los conocimientos actuales referidos a las destrezas escritas y el proceso de aprendizaje se han nutrido de estudios científicos del lenguaje y la psicología de los procesos cognitivos (Hayes, 1996; Hayes & Flower, 1980; Scardamalia & Bereirter, 1992). Aún cuantos estos resultados son numerosos y a veces aparentemente contradictorios, coinciden en destacar la significación del uso y función de la lengua en el proceso de aprendizaje. El papel de la destreza escrita ha evolucionado en los últimos veinte años, considerándose tener hoy un carácter recurrente y cíclico y no ser considerada como antaño un producto lineal (Alcón Soler, 2002: 112).

De un período en el que no se le consideraba desempeñar un papel decisivo (Kahn, 1993), ha adquirido un mayor relieve dentro del proceso de enseñanza-aprendizaje a partir de los lineamientos establecidos por el MCERL y el Portfolio Europeo de las Lenguas⁶³ (Martínez Sánchez, 2004a: 87; Sánchez-Reyes, 2004: 70).

Para el MCERL es ‘texto’ cualquier fragmento de lengua, enunciado o escrito, que los usuarios o alumnos reciben, producen o intercambian en un contexto dado (Instituto Cervantes, 2002:91). Su producción requiere que las competencias lingüísticas se pongan en funcionamiento a través de actividades lingüísticas de comprensión, expresión, interacción y mediación (interpretar o traducir) (Sánchez-Reyes, 2004:7 0).

Martínez Sánchez, en su propuesta didáctica de la destreza escrita en la clase de LE para el enfoque funcional-comunicativo y por tareas, explica que el dominio de la composición pasa por la enseñanza y el aprendizaje de ciertas estrategias cognitivas y lingüísticas del tipo *know how* (saber cómo o dominio del método), donde el aprendiz debe saber cómo reunir, seleccionar y estructurar diferentes tipos de información, saber cómo desarrollar y llevar al texto las ideas y por último saber cómo revisar, valorar y presentar lo escrito (Martínez Sánchez, 2004b: 80); o dicho en otros términos: “Cuando el escritor compone un texto debe tener una visión global y anticipada de lo que quiere decir y cómo lo quiere decir [...] se esfuerza durante el desarrollo de su actividad por organizar sus ideas de modo que el lector pueda comprenderlas [...] hace una revisión para clarificar los significados, reflexionando en torno a lo que ha escrito, releyéndolo y reescribiéndolo para que resulte más comprensible” (Maqueo, 2005: 267-268).

⁶³ Véase en: http://www.coe.int/t/dg4/portfolio/default.asp?l=e&m=/main_pages/welcome.html [08-02-2011].

El alumno que se enfrenta a la tarea de redacción en una lengua extranjera suele disponer de un tiempo de revisión que le falta cuando se expresa de forma oral; sin embargo en el medio escrito no existen los signos extralingüísticos y no verbales (gestos, miradas, etc.), así como las repeticiones y aclaraciones que pueden ser factores decisivos de comprensión en la lengua oral (Carrió Pastor & Seiz Ortiz, 1999: 69).

Además hay diferencias en la forma de redactar entre aprendices noveles y avanzados, donde los primeros escriben textos más cortos, con poca coherencia, omisiones de enlaces y presentación de ideas poco elaboradas, mientras que los más experimentados se concentran en elaborar un texto más rico en contenido y forma debido a que dedican más tiempo a su preparación (Scardamalia & Bereiter, 1992).

Otra cuestión revelada en otros estudios apunta a que los aprendices de LE suelen ser muy restrictivos para expresar puntos de vista e introducir argumentos y debido a esto tienden a utilizar en exceso expresiones y combinaciones de palabras que ya conocen (Granger, 1998: 156).

Otros trabajos investigativos indican que los aprendices de LE al escribir un texto suelen utilizar las mismas estrategias y rutinas que cuando escriben en su LM (Cumming, 1989), de ahí la importancia del proceso de escritura en LM y el conocimiento del profesor de LE de aquellos aspectos lingüísticos que sean similares o diferentes entre las lenguas.

Existen tres dimensiones esenciales del texto: la comunicativa, la pragmática y la estructural (Loureda Lamas, 2009: 20). Estas están a nuestro juicio relacionadas con las competencias comunicativas, a saber: sociolingüísticas, pragmáticas y lingüísticas establecidas en el MCERL (Cap. 5).

De acuerdo con los objetivos del estudio, a fin de correlacionar la destreza escrita/competencia lingüística de los sujetos de nuestro estudio, se han combinado aspectos del listado de descriptores para la evaluación global de las destrezas en la lengua escrita recopiladas en el MCERL (Cap. 4, acápite 4.4.1.2)⁶⁴ y de los

⁶⁴ Véase en: http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cap_04.htm#p42 [26-02-2011].

descriptores por nivel de competencia del usuario contenidos en la propuesta de autoevaluación de González Piñeiro et al. (González Piñeiro et al., 2010: 136).

La correlación entre nivel del aprendiz, descriptores/competencia que aparece en los cuadros 1 y 2, orienta la descripción y evaluación de la competencia de los aprendices y su nivel de destreza escrita que hacemos en el capítulo de discusión (capítulo 7). En primer lugar, aparece una valoración de la comprensión escrita basada en la autoevaluación que los aprendices participantes en la encuesta hicieron de sí mismos; y en segundo lugar se analiza la comprensión escrita contrastando dicha autoevaluación con aspectos de la competencia léxica, sintáctica, gramática y pragmática en la expresión escrita en general, escritura creativa e informes y redacciones, resultado del análisis de la autora del presente estudio.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

Cuadro 1. Descriptores de destreza/competencia lingüística (adaptado de González Piñeiro et al. (2010: 136)

Nivel del aprendiz Escala/MCREL	Descriptores/Destreza
Nivel A2	<p>Comprensión escrita</p> <p><i>Soy capaz de leer textos muy breves y sencillos. Sé encontrar información específica y predecible en escritos sencillos y cotidianos como anuncios publicitarios, prospectos, menús y horarios, y comprendo cartas personales breves y sencillas.</i></p> <p>Expresión escrita</p> <p><i>Soy capaz de escribir notas y mensajes breves y sencillos relativos a mis necesidades inmediatas. Puedo escribir cartas personales muy sencillas; por ejemplo; agradeciendo algo a alguien.</i></p>
Nivel B1	<p>Comprensión escrita</p> <p><i>Comprendo textos redactados en una lengua de uso habitual y cotidiano o relacionado con el trabajo. Comprendo la descripción de acontecimientos, sentimientos y deseos en cartas personales</i></p> <p>Expresión escrita</p> <p><i>Soy capaz de escribir textos sencillos y bien enlazados sobre temas que me son conocidos o de interés personal. Puedo escribir cartas personales que describen experiencias e impresiones.</i></p>

ANÁLISIS CONTRASTIVO Y ENSEÑANZA- APRENDIZAJE DE E/LE EN DOS UNIVERSIDADES SUECAS

Cuadro 2. Correlación entre destreza escrita, nivel y competencias lingüísticas

Destreza escrita	Competencias Lingüísticas
Comprensión escrita	<i>Reflexión del estudiante acerca de esta categoría (encuesta)</i>
Expresión escrita en general	<p><i>Reflexión del aprendiz acerca de esta categoría (encuesta); competencia léxica; competencia sintáctica; competencia gramática; competencia pragmática</i></p> <p><i>Nivel A2: Escribe una serie de frases y oraciones sencillas enlazadas con conectores tales como pero y porque</i></p> <p><i>Nivel B1: Escribe textos sencillos y cohesionados sobre una serie de temas cotidianos dentro de su campo de interés enlazando una serie de distintos elementos breves en una secuencia lineal</i></p>
Escritura creativa	<p><i>Nivel A2: Escribe sobre aspectos cotidianos de su entorno, en oraciones enlazadas; por ejemplo, personas, lugares, una experiencia de estudio o de trabajo. Escribe descripciones muy breves y básicas de hechos, actividades pasadas y experiencias personales. Es capaz de escribir una serie de frases y oraciones sencillas sobre su familia, sus condiciones de vida, sus estudios, su trabajo presente o el último que tuvo. Nivel B1: Escribe relaciones de experiencias describiendo sentimientos y reacciones en textos sencillos y estructurados. Es capaz de escribir una descripción de un hecho determinado, un viaje reciente, real o imaginado. Puede narrar una historia.</i></p>
Informes y redacciones	<p><i>Nivel A2: No hay descriptores disponibles en el MCREL</i></p> <p><i>Nivel B1: Escribe redacciones cortas y sencillas sobre temas de interés. Puede resumir, comunicar y ofrecer su opinión con cierta seguridad sobre hechos concretos relativos a asuntos cotidianos, habituales o no, propios de su especialidad. Es capaz de escribir informes muy breves en formato convencional con información sobre hechos habituales y los motivos de ciertas acciones.</i></p>

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

Los aspectos concretos de la competencia lingüística de los aprendices en cuanto a su expresión escrita aparecen desarrollados en el capítulo siguiente, partiendo de los elementos lingüísticos que se describen y analizan respecto a la IL, o sea las combinaciones de palabras con *que* en los textos escritos por los aprendices suecos de E/LE.

CAPÍTULO 4. COMBINACIONES DE PALABRAS CON *QUE*

So strong are the co-occurrence tendencies of words, word classes, meanings and attitudes that we must widen our horizons and expect the units of meaning to be much more extensive and varied than is seen in a single word (Sinclair, 1996:94)

El propósito principal del capítulo es describir el uso y funciones de las combinaciones de palabras con *que* elegidas en lengua española y sueca. La exposición que sigue constituye la base léxico-gramatical para el posterior análisis de las funciones y frecuencias de uso de estos elementos en textos producidos por aprendices suecos de E/LE (según se desprende del corpus SAELE)⁶⁵. Como analizaremos a continuación, esas múltiples funciones determinan la complejidad de la elección correcta de las combinaciones con *que*. El papel de *que* en español y *att* (conjunción) y *som* (pronombre relativo) en sueco, hace que estas partículas sean en consecuencia los introductores de oraciones subordinadas de mayor frecuencia de uso en ambas lenguas, y que sus combinaciones con otras palabras en contexto incidan en la fluidez de la exposición de los hablantes.

Resulta motivador explorar cómo los nativos suecos aprendices de E/LE usan *que* y sus combinaciones en la producción escrita de la lengua meta debido a que esta partícula en lengua española se corresponde, de acuerdo a las diferentes funciones que cumple, con por lo menos siete otras diferentes en lengua sueca, a saber: *att, som, än, vad, vilken, vilket, vilka*, mientras que la conjunción y relativo *que* en español ha logrado monopolizar funciones muy diversas alcanzando un valor polisémico (Penny, 1993: 148-149, 223).

La elección de una determinada palabra y sus combinaciones en el discurso está determinada en el contexto por el principio de elección libre y el idiomático o fraseológico (Sinclair, 1996: 75-106), formándose un *continuum* entre léxico y sintaxis. Revisar las características del uso de esas palabras y sus combinaciones- según su tendencia idiomática- en ambas lenguas, ofrece la base para la comparación contrastiva español-sueco del uso de las combinaciones de palabras con *que* elegidas: *lo que, de que, algo que, dice que* para detectar posibles similitudes y diferencias en textos

⁶⁵ En este capítulo utilizamos ejemplos de los corpus de control ARGUS (nativos suecos) y CEDEL2 (subcorpus de nativos hispanohablantes).

producidos por nativos suecos estudiantes de E/LE, en comparación con los usos respectivos entre nativos suecos e hispanohablantes.

4.1 Combinaciones de palabras

Las asociaciones repetitivas en el discurso en forma de frases recurrentes (Altenberg, 1998; Altenberg & Eeg -Olofsson 1990; Sinclair, 1991; Stubbs, 1996) han sido denominadas, entre otras, como ‘paquetes léxicos’ (Biber, 2005: 2)⁶⁶, ‘secuencias formulaicas’ (Wray, 2002: 9), ‘patrones prefabricados’ (Erman & Warren, 2000: 31), ‘secuencias convencionales’ (Forsberg, 2010: 29) ‘unidades pluriverbales’ (Moon, 1997: 43), ‘colocaciones’ (Nesselhauf, 2005: 125). Esta diversidad de dominaciones ocurre debido a que estas combinaciones pueden ser largas o cortas, formadas por dos, tres o más palabras, se usan en la lengua con diferentes propósitos y como plantea Schmitt & Carter (2004: 3) pueden ser totalmente fijas o el hablante puede ‘reconstruirlas’ según la situación comunicativa.

Nosotros utilizamos en este estudio la definición de Erman & Warren quienes plantean que estas: *son aquellas combinaciones⁶⁷ formadas por al menos dos palabras que los hablantes nativos prefieren usar en lugar de otra combinación alternativa que podría haber sido equivalente a ésta* (Erman & Warren, 2000: 31).

Estas combinaciones comienzan habitualmente, en lengua hispana, con conjunciones, artículos, pronombres, preposiciones o marcadores del discurso (Butler, 1997: 76). Los estudios de corpus han abordado su presencia en el discurso académico (Biber, 2005; Cortés, 2004; Danielsson, 2007; Simpson, 2004; Verdaguer, 2003, 2004; Wray, 2002).

La elección de combinaciones de palabras en una expresión determinada puede depender del principio de elección libre (*open-choice principle*) que [...] *is the tendency for a word to have a fixed meaning in reference to the world, so that anyone wanting to name its referent*

⁶⁶ El término fue utilizado la primera vez en el *Longman Grammar of Spoken and Written English* (1999) por Biber, Johansson, Leech, Conrad & Finnegan (ibid).

⁶⁷ Pawley & Syder se refieren a estas estructuras como *combinations of word* (Pawley & Syder, 1983: 191-226) y así las llamamos nosotros, aunque utilizamos la definición de Erman & Warren (ellas las denominan *prefabs*/patrones prefabricados).

would have little option but to used it, especially if the relationship works in both directions [...] o del principio idiomático (idiom principle) que [...] is the phraseological tendency, where words tend to go together and make meanings by their combinations. Here is collocation and other features of idiomaticity (Sinclair, 1996: 81-82) pero se ha comprobado que los aprendices de lenguas suelen tener más dificultades al utilizar combinaciones que dependen del idiomático (Flowerdew, 1998, 2003; Granger & Paquot, 2005; Paquot, 2005).

En el proceso de elección de estas combinaciones intervienen mecanismos neurofisiológicos relacionados con el aprendizaje de la lengua materna, la adquisición de segundas lenguas en la infancia y el desarrollo de patrones de fluidez en la lengua oral y escrita (Altenberg, 1998; Altenberg & Eeg –Olofsson, 1990; Nesselhauf, 2005; Sinclair, 1991).

La completa libertad de elección no es lo habitual, libertad y determinismo son dos principios contradictorios que interactúan en la organización de ideas que forman un *continuum* lingüístico debido a que las palabras no se combinan al azar en el discurso oral o escrito sino que el principio idiomático es necesario allí donde el de elección libre no contempla suficientemente la elección consecutiva (Sinclair, 1996: 81; 1991: 110).

Las combinaciones de palabras aparecen como colocaciones de cohesión léxica que relacionan elementos que co-ocurren regularmente (Carter, 1988: 51) y su significado léxico y gramatical en una cadena tiende a cambiar según el contexto de uso (Römer & Schulze, 2009: 2). Es por esta razón que el estudio de las colocaciones tienen un rol central en la enseñanza–aprendizaje de las LE ya que permite descubrir y analizar qué es lo típico en la relación entre palabras en la producción lingüística de un aprendiz (Hersb, 2009:51).

Su uso acertado por aprendices de lenguas es un indicador de las destrezas alcanzadas (Altenberg, 1998; Altenberg & Granger, 2001; Simpson, 2004; Sinclair, 1991; Wray, 2002). Los aprendices de LE en la práctica pueden producir frases diferentes a los nativos (Pawley & Syder, 1983) compuestas en ocasiones por palabras recientemente adquiridas e inapropiadamente empleadas (Foster, 2001), a veces son producto de un intento de ‘acomodar’ sus frases al proceso de comunicación (Seidlhofer, 2009:195-

215, 209, 211)⁶⁸, pero también utilizan colocaciones similares a las elecciones de los nativos (Siyanova & Schmitt, 2008).

La selección de estas combinaciones para un hablante nativo equivale a la emisión de una unidad léxica completa con su significado y forma fonológica asociada, mientras que el estudiante de lengua debe en ocasiones reconstruirla a partir de la elección de sinónimos y las posibles interferencias de su LM (Wray, 2005: 24). La memorización de combinaciones de palabras en la lengua meta es decisiva en el desarrollo de niveles superiores de fluidez comparables a las de los nativos (Aijmer, 2002; Hasselgren, 2002; Simpson, 2004; Wray, 1999, 2002).

Entre los estudios publicados acerca del uso de combinaciones de palabras aparecen varios realizados en universidades suecas, que se sitúan en diferentes ramas lingüísticas, pueden tener muestras de lengua oral o escrita y analizan el fenómeno desde diferentes perspectivas. Solamente encontramos uno referido a aprendices suecos de E/LE, pero se diferencia del nuestro en que en el análisis se parte desde la perspectiva de interacción en la conversación, se basa en muestras de la lengua oral, aparte de que las herramientas metodológicas empleadas no son las mismas (Moreno Teva, 2007).

Además, encontramos otros estudios en que se utilizaban la LC y el análisis contrastivo con el objetivo de indagar, por ejemplo, el nivel de destreza en lengua oral y escrita entre nativos y no nativos, aunque los participantes eran aprendices suecos de francés e inglés. En los estudios contrastivos sueco-francés destaca el de Forsberg (2010), quien realiza una descripción del uso oral de secuencias convencionales entre participantes de francés L2 y nativos de francés. Los resultados muestran que las secuencias convencionales léxicas son las más difíciles de adquirir por los nativos suecos de francés L2, dado que solamente los aprendices más avanzados obtuvieron resultados más cercanos a los nativos.

En el caso de aprendices suecos de inglés el estudio más relevante para nosotros es el de Wiktorsson (2003), dado que la autora investiga la lengua escrita de aprendices suecos de inglés en diferentes niveles de aprendizaje y compara los resultados con

⁶⁸ Según estudios realizados sobre el inglés como lingua franca. El término lingua franca se utiliza para referirse a una lengua que sirve como medio de comunicación entre los miembros de una comunidad lingüística cuyas LM son diferentes. Por ejemplo: el Swahili en la parte del este y central de África o el inglés en el mundo de hoy.

nativos anglohablantes. Su estudio muestra que los aprendices suecos más avanzados en lengua inglesa no usan significativamente un menor número de unidades pluriverbales que los nativos, y que presumiblemente se debe a que utilizan combinaciones más propias del registro oral. Por otra parte, los aprendices menos avanzados utilizan estrategias basadas en la lengua materna por lo que construyen frases que carecen de idiomática, al parecer debidas a traducciones literales del sueco.

4.2 Estudios precedentes de combinaciones de palabras con *que* entre nativos suecos aprendices de E/LE

En dos estudios anteriores habíamos analizado similitudes y variaciones del uso de combinaciones de palabras con *que* entre aprendices suecos de E/LE (Pino, 2006; Pino, 2010). En el primer estudio se compararon 27 resúmenes de textos escritos por aprendices suecos (nivel B1-B2) con dos corpus de referencia de nativos hispanohablantes⁶⁹, utilizando el procesador textual AntConc 3.1.302⁷⁰. El resultado mostró que la partícula *que* era la más frecuentemente empleada en los textos de los aprendices suecos, apareciendo 67 veces cada 1 000 palabras, mientras que en el Corpus del español apareció 21 veces cada 1 000 palabras de texto y 35 en el Corpus SOL- Spanish Online. Las secuencias de palabras en asociación con *que* más utilizadas fueron: *lo que*, *dice que*, *de que*, y *algo que*. *Lo que* fue la combinación más frecuentemente empleada tanto por suecos aprendices de E/LE como por nativos hispanohablantes de los corpus de referencia. Sin embargo su frecuencia de uso por los suecos cuadruplicó a la de nativos hispanohablantes. La combinación *dice que* se utilizó 28 veces más frecuentemente por los aprendices suecos que en los corpus de referencia, y *algo que* 12 veces más frecuentemente. El resultado mostró a la partícula *que* como el recurso idiomático más utilizado por los aprendices suecos de E/LE para la construcción de combinaciones de palabras en comparación a los aparecidos en el Corpus del español⁷¹ y en el corpus SOL- Spanish Online (Pino, 2006).

⁶⁹ Corpus del español y el Corpus SOL- Spanish Online.

⁷⁰ El AntConc 3.1.302 tiene herramientas similares al WordSmith.

⁷¹ Davies, Mark. (2002-) Corpus del Español (100 million words, 1200s-1900s). Available online at <http://www.Corpusdelespanol.org>.

En el segundo estudio analizamos similitudes y variaciones en el uso de combinaciones de palabras con *que*, en 30 textos producidos por 10 aprendices suecos de E/LE (nivel A2-B1), que fueron comparados con los resultados obtenidos del corpus de aprendices suecos (nivel B1-B2) utilizado en el estudio anterior y con el Corpus del español. Los textos fueron procesados con las herramientas del WordSmith. Los resultados mostraron la frecuencia de *que* 20/1 000 en el Corpus SAELE, 67/1 000 en el corpus de aprendices (nivel B1-B2) y de 21/1 000 en el Corpus del español. La combinación *lo que* fue la más frecuentemente empleada en los tres corpus, pero su frecuencia de uso por los aprendices suecos fue casi cuatro veces mayor que la de los hispanohablantes. El uso de las combinaciones *de que*, *algo que* y *dice que* disminuyó significativamente entre los aprendices de E/LE en el nivel B1-B2 en comparación con los del nivel A2-B1, pero se mantuvieron siempre en cifras al menos del doble de frecuencia de uso en comparación con el Corpus del español (nativos hispanohablantes).

La combinación *lo que* fue la más frecuentemente empleada en los corpus utilizados, el de aprendices suecos de E/LE, nivel A2-B1, el de aprendices suecos de E/LE, nivel B1-B2, y el corpus de nativos hispanohablantes (Corpus del español), aunque su frecuencia de uso por los aprendices suecos fue casi cuatro veces mayor que la de nativos hispanohablantes (Pino, 2010).

4.3 Combinaciones de palabras con *que*

Se resumen a continuación las funciones principales de *que* en español⁷² y de *att*, *som*, *än*, *vad*, *vilken*, *vilket* y *vilka* en sueco, con el propósito de guiar el análisis posterior de forma más coherente para el lector.

⁷² Para el resumen del cuadro 3 hemos utilizado el *Diccionario de uso del español*, el *Norstedts svenska ordbok* / Diccionario de sueco y el *Norstedts spansk-svenska, svensk-spanska ordbok* / Diccionario español-sueco y sueco español. Para el análisis también nos basamos en: la *Nueva gramática de la lengua española, Gramática descriptiva contemporánea de la lengua española, Gramática de la lengua española (de Alarcos Llorach), Gramática española por niveles* y *A New Reference Grammar of Modern Spanish*. Precisamos que para el análisis contrastivo español-sueco hemos además utilizado *Spansk grammatik/ Gramática española, Svenska Academiens grammatik (Gramática de la academia sueca)* y *Descriptiv svensk grammatik/ Gramática descriptiva sueca*. Es bueno aclarar que por limitaciones del estudio es imposible abarcar todas las funciones de *que* por lo que presentamos solamente las más significativas a partir de su relación con las combinaciones de palabras que estudiamos.

COMBINACIONES DE PALABRAS CON *QUE*

Cuadro 3. *Que/ att, som, än, vad, vilken, vilket y vilka*

QUE EN ESPAÑOL	ATT, SOM, ÄN, VAD, VILKEN, VILKET, VILKA EN SUECO
<p>Que>Conjunción subordinante y completiva= introduciendo subordinadas</p> <p>-obligación>con el verbo tener</p> <p>-conjunción optativa y coordinante en expresiones fijas</p> <p>-conjunción causal = porque</p> <p>-consecutiva</p> <p>-comparativa</p> <p>-locuciones conjuntivas</p>	<p><i>Att</i>> conjunción, introductor de oraciones subordinadas</p> <p><i>Att</i>> marca de infinitivo e introductor en oraciones exclamativas que no tiene correspondencia en español</p> <p><i>Som</i>= como> conjunción, introductor de oraciones subordinada y comparativa=<i>såsom</i></p> <p><i>Mer/ mindre... än</i>> conjunción comparativa</p> <p><i>Än</i>>introductor de oraciones subordinadas</p>
<p>Que>Pronombre relativo, introductor en oraciones del relativo</p> <p>Con artículo determinado equivale a ‘el cual’ y sin antecedente y referido a persona a ‘quien’</p> <p>Especificativa</p> <p>Explicativa</p> <p>Valor ponderativo</p>	<p><i>Som</i>> pronombre relativo, introductor en oraciones del relativo</p> <p><i>Vilken/ vilket/ vilka</i>>pronombres relativos</p> <p><i>Som, vilken/ vilket/ vilka</i>>corresponden a que, el que, quien y el cual en español</p>
<p>¿Qué?>Pronombre interrogativo</p>	<p><i>Vad, vad för/ en, vad för slags, vad för någon(ting) vilken/ vilket/ vilka</i></p>
<p>¡Qué!>Pronombre exclamativo</p>	<p><i>Vad, vilken/ vilket/ vilka, så</i></p>

En las traducciones español-sueco suelen emplearse diferentes palabras según el contexto, de manera que a *que* puede corresponderle *att, som, vilken/vilket/vilka, vad, än*, de la misma forma que en las traducciones sueco-español el pronombre relativo sueco *som* podría sustituirse por *que, el que, quien, el cual* (Benson et al., 1999; Fält, 2000: 164). En sueco, las palabras equivalentes a *que* (*att, som, vilken/vilket vilka, vad,*

än) son registradas por *Svenska Akademiens grammatik* en funciones conjuntiva, relativa, interrogativa-exclamativa y comparativa que se emplean en diferentes contextos en combinación con palabras adyacentes (Teleman, Hellberg & Andersson, 1999).

Las conjunciones constituyen una clase de palabras invariables, generalmente átonas que relacionan entres sí vocablos y grupos sintácticos, unas veces equiparándolos y otras jerarquizándolos o haciéndolos depender unos de otros, donde están las coordinantes (las que enlazan) y las subordinantes que los vinculan marcando entre éstas relaciones de dependencia (NGLESII⁷³, §31.1a, 2009: 2395).

La palabra *que* como conjunción subordinante completiva y final convierte a la subordinada que encabeza en sustantiva que actúa frecuentemente como objeto directo o sujeto de diversos tipos de verbos (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008: 825), como en los ejemplos:

(1) Es simplemente deleznable *que* nos hagan aprender, respirar y fumar

(subcorpus CEDEL2)

Tener +*que* (obligación)

(2) Al fin llegamos al hospital y tuvimos *que* esperar una eternidad

(subcorpus CEDEL2)

En expresiones muy fijas

(3) Está el día que llueve, *que* no llueve= ya llueve ya no llueve

(Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008: 830)

Cuando se establecen unidades fuertes entre preposición y adverbio, se forman conjunciones compuestas como la causal *porque*, la concesiva *aunque* y la final *para que* (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008: 827). La conjunción *que* puede sustituir en contexto algunas de estas conjunciones como en los ejemplos siguientes donde desempeña función causal:

⁷³ Utilizaremos la sigla NGLESII para referirnos a Nueva gramática de la lengua española, Sintaxis II.

(4) Me llamó la atención la forma de vestir de la gente *que* (porque) eran especialmente elegantes

(subcorpus CEDEL2)

(5) La veracidad de la historia ayuda mucho *que* (para que) te identifiques en esta historia

(subcorpus CEDEL2)

Comparativa en:

(6)...en comparación llueve mucho más *que* en Sevilla

(subcorpus CEDEL2)

La conjunción *que* también forma parte de un gran número de locuciones conjuntivas en las que *que* se podría combinar con otras estructuras teniendo la siguiente forma (NGLESII, 2009: 2459-63) en:

(7) Quedaban cinco minutos *para que* saliera mi tren (preposición para + que)

(subcorpus CEDEL2)

De la misma forma que *que* conjunción posee múltiples valores, *att* y *som* (unos de sus equivalentes en lengua sueca) tienen un significado muy general y uso variado (Teleman, Hellberg & Andersson, 1999: 734). En lengua sueca se clasifican las palabras *att* y *som* como conjunciones subordinantes/*subjunktioner*, lo que significa que forman oraciones que son complementos de otras y son palabras introductoras en la mayoría de las oraciones subordinadas [...] siendo parte de frases comparativas, predicativas y concesivas (Teleman, Hellberg & Andersson, 1999: 733) como por ejemplo *att* en una predicativa:

(8) Jag vet *att* han inte kommer i dag heller/Sé que él tampoco viene hoy

(Holm & Nylund, 1979: 197)

Att introduce oraciones narrativas y consecutivas, así como un tipo expresivo de oración principal, aparte de cumplir una función como marca de infinitivo cuando comienza frases de este tipo, mientras que *som* introduce oraciones subordinadas relativas, expresivas, comparativas, temporales y causales, así como también oraciones principales expresivas y comparativas (Teleman, Hellberg & Andersson, 1999: 734,745).

Desde el punto de vista sintáctico *att* puede iniciar la oración subordinada o la oración principal, por ejemplo:

(9a) De meddelade *att* han just lagat datorn/Dijeron *que* acaba de arreglar la computadora

(Teleman, Hellberg & Andersson, 1999: 743)

(9b) *Att* du bara vill vara med om det/Que sólo quieras ser parte de ello

(Teleman, Hellberg & Andersson, 1999: 734)

También suelen formarse combinaciones como: preposición + *att* {*för att*/porque, *utan att*/sin que} (Teleman, Hellberg & Andersson, 1999: 739) como vemos en:

(10) Språket i en barnbok är också viktigt *för att* utvidga barnets ordförråd/El idioma en un libro infantil es también importante *para* enriquecer el vocabulario del niño

(subcorpus ARGUS)

El relativo *que* puede actuar como nexo de subordinación introduciendo una cláusula como oración dependiente desempeñando una función sintáctica en la subordinada (sujeto, objeto directo, etc.), o con una naturaleza anafórica (que permite interpretar semánticamente la subordinada en relación con el grupo nominal del que forma parte). En el ejemplo: el libro que leo>, en el cual *que* cumple las tres funciones al ser un nexo subordinante, el complemento directo de> leo y un elemento anafórico cuyo antecedente es>libro (NGLESII, 2009: 3297; Martí & Torrens, 2001: 49).

Que es el pronombre relativo de mayor uso (Butt & Benjamin, 2000: 494). Con artículo determinado equivale a *el cual*:

(11) Y no despreciamos el idioma con el que/el cual ahora escribo

(subcorpus CEDEL2)

Sin antecedente y referido a persona equivale a *quien*:

(12) Trata sobre un policía californiano *que*/quien recibe un día una carta

(subcorpus CEDEL2)

Tiene además las siguientes funciones, según el *Diccionario de uso del español* (Jarraud, Ramón, S., Ramón, F. & Ramón, H., 2008: 1396-1397):

Especificativa:

(13) Cuenta la historia de una chica *que* vive en un barrio de clase humilde en Inglaterra

(subcorpus CEDEL2)

Explicativa:

(14) la vida cotidiana de una tribu indígena *que*, aislada del resto de las tribus, vive amenazada por otra tribu

(subcorpus CEDEL2)

Posee valor ponderativo en frases en que el antecedente es un adjetivo, un nombre de cualidad sin artículo o una expresión de situación o actitud:

(15) Pues eso, *que* para mí tienen todo el derecho

(subcorpus CEDEL2)

Las funciones de *que* como conjunción o relativo no son siempre evidentes, debido a que los términos adyacentes pueden influir en su significación (Alarcos, 1994: 260), pudiendo aparecer en función de conector subordinante (Martí & Torrens, 2001: 7-8) como en el caso de su papel completivo en las oraciones subordinadas sustantivas, por ejemplo:

(16) pues puedo decir *que* he visto casi todo

(subcorpus CEDEL2)

En ocasiones forman *que* y *lo* un morfema (locución, conjunción subordinante o coordinante), que enlaza dos o más enunciados con fines argumentativos únicos (Escandell, 2008: 98-100), como una conexión entre dos o más oraciones, cada una con función argumentativa específica y ambas forman un bloque argumentativo (Loureda & Acín, 2010: 25). Estos bloques parecen ser más típicos de la lengua escrita (Domínguez, 2007; López & Borreguero, 2010: 420, 463) por su nivel de integración sintáctica que es superior al de la lengua oral (Montolío, 2001: 35-41).

En la lengua sueca *som* es el pronombre relativo de uso más común y puede referirse a personas, animales o cosas tanto en singular como en plural. Puede aparecer como sujeto u objeto en la oración. En ocasiones puede ser sustituido por los relativos: *vilken*, *vilket* y *vilka* (Holm et al., 1979: 66). Cuando no cumple el papel de sujeto puede omitirse: [*Boken (som) jag läste igår*/El libro que leí ayer]. En este ejemplo se ve claramente que tal omisión no es válida en español: El libro \emptyset leí ayer.

Como relativo:

(17) Det kanske finns lagar som ska skydda henne/Quizás existan leyes que la protejan

(subcorpus ARGUS)

(18) Varför jag valde denna bok *som* jag inte minns namnet på/Por qué elegí éste libro del que no recuerdo el nombre

(subcorpus ARGUS)

La lengua sueca no admite preposición antes del pronombre relativo *som*, como se aprecia en el ejemplo con la preposición *på*/del. Obsérvese además que en este caso *som* podría ser equivalente a *vilken*.

De los pronombres relativos suecos, solamente *vilken* tiene formas para el singular y el plural, que serían *vilken* (*utrum singularis*), *vilket* (*neutrum singularis*) y *vilka* (*pluralis*), además de que cuando es un atributo posesivo se utiliza la forma del genitivo: *vilkens*, *vilkets* y *vilkas*. *Vilken* se utiliza más en la lengua escrita formal, y puede corresponderse con *som* o *något som* en los usos más coloquiales (Teleman, Hellberg & Andersson, 1999: 349).

Vad es un pronombre interrogativo-exclamativo en sueco, que se usa además en preguntas indirectas y en construcciones combinado con *som* (Allén et al. 1990: 1122). Como introductor en oraciones subordinadas relativas es equivalente a *det som/lo que* (Wijk-Andersson, 1985: 36).

El uso de *que* en función interrogativa:

(19) ¿*Qué* sucedía cuando...

O exclamativa:

(20) ¡*Qué* bueno saber de ti!

(subcorpus CEDEL2)

El *que* interrogativo-exclamativo se corresponde con *vad* (21 y 22). También hay un ejemplo en que *vad* aparece en una pregunta indirecta combinado con *som* (23).

(21) *Vad* händer om vi inte läser?/ ¿*Qué* sucede si no leemos?

(22) *Vad* härligt att det finns böcker!/ ¡*Qué* maravilloso que hay libros!

(23) Genom att läsa får barnet veta *vad som* är rätt och fel, sant eller falsk/ A través de la lectura el niño puede saber *qué / lo que* es correcto e incorrecto, verdad o mentira

(subcorpus ARGUS)

El cuadro 4 presenta la correspondencia en lengua sueca de las combinaciones con *que*: *lo que*, *de que*, *algo que* y *dice que* y servirá de base para el análisis contrastivo de sus

formas y frecuencias de usos entre nativos suecos aprendices de E/LE en comparación con nativos hispanohablantes y suecos, respectivamente, de los corpus de control correspondientes.

Cuadro 4. Correspondencia de combinaciones de palabras con *que*: *lo que*, *de que*, *algo que* y *dice que*

COMBINACIONES EN ESPAÑOL	COMBINACIONES EN SUECO
Lo que	<i>vilket/vilken/vilka; något som, vad som, det som/ den som; så</i>
De que	<i>att, vad</i>
Algo que	<i>något som, vilket, vad, vad som</i>
Dice que	<i>Säger att/ säga åt något= /le/dice que</i>

4.3.1 Combinación *lo que*

El relativo *que* ha de ir precedido del artículo cuando no hay un antecedente expreso: {Puedes escoger *la que* quieras}; en las oraciones separadas con coma que sirven para especificar el antecedente: {Dame esa hoja, *la que* está en blanco}; en las construcciones enfáticas con el verbo ser: {Siempre es él *el que* escoge primero} y tras preposición: {El coche con *el que* hicieron el viaje tenía más de veinte años} (Martí & Torrens, 2001: 23).

Lo que puede cumplir la función de conector textual cuando tiene como antecedente a una oración (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008:883). Se utiliza ante infinitivo en preguntas indirectas como *que* (Butt & Benjamin, 2000: 341).

En el corpus de control CEDEL2 (subcorpus de nativos hispanohablantes) aparecen además varias combinaciones de palabras con *lo que* en expresiones como: *dar por hecho lo que no está hecho; ver más allá de lo que pueden ver y hacer lo que quiera*.

Una combinación muy usual es: preposición + *lo que* donde las preposiciones predominantes son: por + *lo que*; con + *lo que*; en + *lo que* y de + *lo que*.

Otra combinación de palabras es verbo + *lo que*; creer + *lo que*; ser + *lo que* y o sería que + *lo que*.

Dado que la cláusula no siempre es restrictiva es común que se use *lo cual* en lugar de *lo que* en el discurso escrito (Butt & Benjamin, 2000: 500).

Lo que se traduciría al sueco dependiendo del contexto como: *vilken, vilken, något som, vad som, det som* y *så* (Benson, Strandvik & Sánchez, 1999; Fält, 2000: 171) y esto se observa en el subcorpus de control ARGUS.

(24) Barnen har omänskliga arbetstider *vilket* medför dåligt dygnsvila/Los niños tienen horas de trabajo inhumanas *lo que* trae consigo poco descanso

(25) Följande problem är utbildning *vilket* många aldrig får/El siguiente problema es la educación a *la que* muchos de ellos nunca tienen acceso

(subcorpus ARGUS)

En estos dos casos se utiliza el equivalente *vilket* y esto depende de la relación entre el pronombre y el sustantivo en lengua sueca. Aquí se refiere a la relación pronombre-*ett*-sustantivo o frase, lo cual determina que el pronombre sustituyente de *som* sea *vilket* para un *ett* sustantivo, *vilken* para un *en* sustantivo o *vilka* para el plural. A diferencia de *som*, '*vilket, vilken* y *vilka*' pueden tener como antecedente a una preposición (Wijk-Andersson, 1985: 35). El género del sustantivo en sueco puede ser *utrum*- masculino, femenino y *realis* (=no animado ni neutro) och *neutrum* y pueden estar en singular o plural, en forma determinada o indeterminada así como estar en forma genitiva o básica (Sörhman, 2007: 46; Teleman, Hellberg & Andersson, 1999: 62).

(26) Ungdomar duar varandra, "ni" är *något som* används för att visa respekt/Los adolescentes se tutean entre sí,"usted" es *algo que/lo que* se utiliza para mostrar respeto

(subcorpus ARGUS)

En este caso se utiliza *något som* con sus posibles equivalentes al español *algo que* y *lo que*. La combinación es elegida por su función de aclaración determinada por el contexto de la oración.

(27) Detta är precis *vad som* händer ute i världen/Esto es precisamente *lo que* sucede en el mundo

(subcorpus ARGUS)

Aquí también se analiza el uso de la combinación a partir de su relación con las otras palabras en contexto donde además está determinado el uso de la combinación por el carácter enfático que se pone en la frase.

(28) Och det var ju *det som* gjorde böckerna så bra/Y eso fue precisamente *lo que* hizo que los libros fuesen tan buenos

(subcorpus ARGUS)

En el ejemplo se utiliza la combinación *det som* con un valor especificativo y cierto énfasis ‘fue eso precisamente’ (y no otra cosa) lo que le otorgaba mérito a los libros en cuestión. En este último ejemplo vemos como *som* está en relación directa con ‘de civila människor’/los civiles.

(29) I krig så är det de civila människorna som är drabbade/En la guerra son los civiles *los que* se ven afectados

(subcorpus ARGUS)

4.3.2 Combinación *de que*

La combinación *de que* puede tener los siguientes equivalentes suecos: *att, vad* (Benson, Strandvik & Sánchez, 1999). *De que* se utiliza como introductor de oraciones subordinadas cuando *que* es una conjunción y no un pronombre relativo, aparece en combinaciones con verbos que denotan un estado mental o de emoción (Butt & Benjamin, 2000: 446-447), como vemos en:

(30) porque temía *aburrirme /de/ que* el viaje fuese muy largo

(31) Se indignaba */de/ que* la situación no cambiara

(subcorpus CEDEL2)

Suele ir precedido de una preposición cuando el verbo principal lo exige en las subordinadas sustantivas (Martí & Torrens, 2001: 19) y además si forma un complemento con ésta (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008: 826).

La combinación *de que* está en relación con oraciones subordinadas relativas que expresan causa, motivo o razón y cuyo verbo debe estar en subjuntivo (Fält, 2008: 340). Por ejemplo:

(32) Me parece buena la idea *de que* no se pueda fumar

(subcorpus CEDEL2)

En el ejemplo 32 observamos primero: me parece buena la idea (razón) + *de que* + pueda.

Que también puede llevar adelante la combinación *el hecho de* cuando la subordinada actúa como sujeto, y en este caso el verbo va en subjuntivo (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008:826), como en el ejemplo:

(33) *El hecho de que* exista me sorprende

(subcorpus CEDEL2)

No obstante, el hablante también podría decir *el hecho de que existe*, con el verbo en modo indicativo, y esta elección depende del conocimiento que se supone tengan los hablantes de la cuestión que se trate (Fält, 2000: 350). Otro caso en el que suele utilizarse el subjuntivo es cuando la oración comienza con *el hecho de que*, aunque la combinación puede aparecer también casi al final, por ejemplo:

(34) *El hecho de que* en España se haya dado un paso

(35) y no lo quieran ver por el simple *hecho de que* seamos catalanes

(subcorpus CEDEL2)

Otras combinaciones de palabras con *de que* forman parte de oraciones subordinadas que pueden o no requerir el modo subjuntivo. Este modo se usa en lengua sueca de forma opcional cuando se refieren a un hecho cuya acción se presupone que sea irreal o imposible, y actualmente el uso es raro excepto la forma *vore (de vara)/ sería (de ser)* (Teleman, Hellberg & Andersson, 1999: 549), mientras que el subjuntivo en la lengua española tiene un amplio uso lo que hace que el verbo de la subordinada vaya en subjuntivo cuando el verbo de la oración principal expresa deseo, influencia, posibilidad y necesidad y los llamados verbos factitivos como conseguir, contribuir, dejar, disuadir, evitar, hacer, lograr, obtener[...] (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008: 689-690).

A continuación analizamos algunas especificidades de uso de la combinación *de que*:

- Se utiliza con subordinadores de tiempo: *antes (de) que/inman* y *después (de) que/efter det att/sedan* (Butt & Benjamin, 2000: 265, 266; Fält, 2000: 352), en las variantes del español de América se utiliza el indicativo en la última combinación.

(36) al rescate de la chica, justo *antes de que* el coche la atropellase

(subcorpus CEDEL2)

- Se emplea con subordinadores de condición: con tal *(de) que*; a condición *de que*; con la condición *de que*; bajo (la) condición *de que*; a cambio *de que* (Butt & Benjamin, 2000: 453).

(37) Te llevo *a cambio de que* me des un beso

(subcorpus CEDEL2)

- Precedido de verbos como desear, querer o adverbios como necesidad, posibilidad: deseo *de que*, la necesidad *de que*, la posibilidad *de que* (Fält, 2000: 351).

(38) Mis sueños estaban sujetos *a la posibilidad de que* la vida me castigara

(subcorpus CEDEL2)

- En el llamado dequeísmo que ocurre cuando se antepone la preposición *de* a una oración subordinada sustantiva de complemento directo como por ejemplo en: {Me dijeron *de que* se iban a cambiar de casa} correcto: {Me dijeron *que* se iban a cambiar de casa⁷⁴}. Se da con mayor frecuencia con los verbos de comunicación y creencia (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008: 698).

- En conjunciones finales como: *para que*, *a fin de que* y con (el) objeto *de que* (Fält, 2000: 354).

(39) Det är viktigt för barn att läsa *för att* utveckla sitt språk /Es importante para el niño leer *para que/a fin de que/con el objeto de que* desarrolle su idioma

(subcorpus ARGUS)

En los ejemplos del corpus Argus a continuación veremos cómo se utilizan *vad* y *att* como equivalentes de *de que*.

(40) Guillou vet inte *vad* han talar om/Guillou no sabe *de qué/de lo que/qué* está hablando

(41) Risken för *att* föräldrar har allt mindre tid för sina barn/El riesgo */de/ que* los padres tienen cada vez menos tiempo para sus hijos

⁷⁴ En: <http://buscon.rae.es/dpdI/SrvltConsulta?lema=dequeismo>

4.3.3 Combinación *algo que*

Algo se utiliza para indicar la existencia de una cosa en relación con lo que se dice en el enunciado del que forma parte (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008: 282). El pronombre indefinido/adverbio *algo* corresponde a *något* y *någonting* en sueco (Fält, 2000: 176). El pronombre *algo* puede aparecer seguido de *que* y en este caso los equivalentes en sueco de *algo que* son las combinaciones *något som*, *vilket*, *vad*, y *vad som* (Benson, Strandvik & Sánchez, 1999).

(42) Sexualitet är *något som* alla på *något sätt* upplever/La sexualidad es *algo que* todo el mundo experimenta de alguna forma

(43) Han har höga positioner på arbetsplatsen *vilket* leder till att arbetet tar mycket tid/Él tiene una alta posición en su trabajo, *algo que resulta en/lo que hace/que* el trabajo le tome/lleve mucho tiempo

(subcorpus ARGUS)

En los dos ejemplos el uso está determinado por la relación sustantivo-congruencia en la frase. En el primer caso está el pronombre *algo/något* directamente en relación con un sustantivo *ett* (*sexualite-t*) y en el segundo *vilket* porque se refiere a toda la frase como explicamos más abajo.

Algo que se puede utilizar en lugar de *lo que* cuando hace alusión a toda la oración y entonces se corresponde con *vilket* y *något som* en lengua sueca (Fält, 2000: 171). La construcción *algo que* + infinitivo es una de las subordinadas de relativo con infinitivo donde el antecedente, en nuestro caso *algo* es un cuantificador pronominal no específico. Véase en:

(44) ¿Tienes *algo que* objetar?

(Bosque & Demonte, 2000, Tomo II: 2290)

La combinaciones *ser* + *algo que* aparece en el corpus CEDEL2 (hispanohablantes).

(45) El problema del terrorismo es *algo que* a los españoles nos afecta

(subcorpus CEDEL2)

(46) *fantasin är något som* alla använder mycket ofta/la fantasía es *algo que* todos utilizan muy frecuentemente

(subcorpus ARGUS)

En relación con esta combinación *algo que* en contexto con la preposición *con* se han observado errores en estudios preliminares acerca de los aprendices suecos de E/LE (Pino, 2006, 2010) del que seleccionamos un ejemplo.

(47) porque tienen que ofrecer al lector *algo que* el pueda identificarse *con*/Eftersom de måste erbjuda läsaren något som han kan identifiera sig *med*

En este caso el aprendiz coloca la preposición *con* al final de la frase tal y como se hace en sueco, en inglés y otras lenguas germanas en que hay diferencias respecto al orden de las palabras y la construcción (Sörhman, 2007: 69). Si tomamos un ejemplo del corpus ARGUS, podemos ver también cómo los nativos suecos utilizan la preposición *till* (en el ejemplo 48) y *för* (en el ejemplo 49).

(48) Att få dem att känna att det är något roligt och inspirerande, inte *något som* man tvingas *till*/Para hacerles sentir que es algo divertido y que inspira y no algo a lo que lo obligan a uno

(subcorpus ARGUS)

(49) Böcker var inte alls något som man intresserade sig *för*, utan man började att "hänga" på stan/Los libros no eran para nada algo que interesara, sino que empezamos a andar/deambular/'mataperrear' por la ciudad

(subcorpus ARGUS)

4.3.4 Combinación *dice que*

La combinación *dice que* corresponde en sueco a *säger att*/'decir que alguien hace algo' o *säga åt något*/'decir a alguien que haga algo' (Benson, Strandvik & Sánchez, 1999; Fält, 2000: 339). En lengua sueca *att* se utiliza con frecuencia con verbos de comunicación (*sägaverb* en sueco) y con el tiempo ha ocupado una posición entre conjunción y signo de puntuación/*skiljetecken* (Holm, Nylund et. al., 1979: 156).

En la lengua española el verbo *decir* es el representante del acto del habla por excelencia [...] y es uno de los verbos con que más comúnmente se emplea el dequeísmo (DeMello, 1994: 181).

Decir como verbo de comunicación puede expresar un hecho (indicativo) o un mandato (subjuntivo) [...] cuando estos verbos introducen un hecho algo real, ya

observado, es normal una cláusula completa con el indicativo, pero esta puede reducirse (sobre todo en estilos literarios y periodísticos) con la identidad de sujetos [...] y esta reducción no es posible cuando el verbo introduce un mandato o deseo (subjuntivo) (Stanley & González, 2007: 235).

Otro aspecto relacionado con el uso de los verbos de comunicación (entre ellos *decir*) y *que* está en relación con el estilo cuando por ejemplo, en el estilo indirecto se convierte el mensaje original en una oración subordinada sustantiva de objeto directo introducida por *que* y dependiente de un verbo de comunicación [...] y cuando el cambio del estilo de directo a indirecto lleva consigo la sustitución del imperativo por subjuntivo (Martí Sánchez, M., Penadés Martínez, I. & A. M. Ruiz Martínez, 2008: 694-695).

Decir que puede ir seguido de un verbo en infinitivo, en indicativo o en subjuntivo (Fält, 2000: 339, 445) y la elección de uno u otro modo suele causar dificultades a los aprendices suecos de E/LE.

Decir que requiere del subjuntivo después de *que* cuando el sujeto en la oración principal no corresponde al de la subordinada (Butt & Benjamin, 2000: 246) y cuando *decir* significa en sueco *säga åt något* 'decir algo a alguien' o sea cuando pasa a ser un verbo de influencia (Fält, 2000: 405-406). El verbo de comunicación (decir) pasa a ser interpretado como un verbo de influencia, experimentando un proceso de reinterpretación semántica que induce el subjuntivo en el verbo de la cláusula subordinada, puesto que se desconoce si se realizará la acción (Lozano González, 2005).

Compárese los siguientes ejemplos donde aparece *decir que* seguido del verbo cantar en infinitivo (ejemplo 50), *decir que* como verbo de influencia seguido del verbo ir en subjuntivo (ejemplo 51) y en indicativo (ejemplo 52).

(50) Puedo *decir que* cantar es mi gran pasión (yo + *decir que* + infinitivo)

(51) Le puedes *decir que* vaya a ver a su hermano (yo>te 'mando a *decir*' + *que* + subjuntivo)

(52) Carmen *dice que* viene temprano (ella + *dice que* + presente indicativo >acción de futuro)

Los aspectos lingüísticos descritos aquí sobre *que* y las combinaciones de palabras: *lo que*, *de que*, *algo que* y *dice que* junto a los resultados que presentamos en el capítulo

COMBINACIONES DE PALABRAS CON *QUE*

posterior (6) los utilizaremos en el análisis y discusión del capítulo 7 para contestar a las preguntas de investigación de este estudio.

CAPÍTULO 5. MATERIAL DE TRABAJO

En el presente capítulo se discute la recopilación y elaboración del corpus SAELE (Suecos aprendices de español lengua extranjera), y de los subcorpus de control ARGUS (nativos suecos) y CEDEL2 (nativos hispanohablantes) (5.1-5.2.2). Además se presenta y analiza el material de trabajo utilizado para recoger los datos de los participantes (encuesta), describiendo especialmente las categorías elegidas en su diseño (5.3-5.3.1).

5.1 El corpus SAELE

En la conformación del Corpus SAELE seguimos fundamentalmente las recomendaciones metodológicas de Sinclair (1991, 1996, 2005), Biber (1993) y el modelo de Lozano (2008). Asimismo nos guiamos por aspectos generales propuestos para elaborar un corpus lingüístico (Lüdeling et al. 2008; Hunston, 2006; Flowerdew, 2008; Torruella & Llisterri, 1998).

Para la definición de los criterios de inclusión/exclusión en la selección de los sujetos, la delimitación del objeto de estudio (tipos y cantidad de textos), y aspectos del método (principalmente herramienta para el procesamiento de los textos), han sido tomados como referencia el Corpus ICLE, el ARGUS y el CEDEL2. Las características de esos corpus las hemos analizado para adaptarlas al propósito específico del presente estudio, al contexto propio del universo y la muestra del mismo y a las particularidades de los sujetos elegidos para la elaboración del corpus SAELE.

Como apuntamos en el capítulo 2, al explicar los pasos seguidos en la elaboración del corpus SAELE nos guiamos por diez recomendaciones metodológicas que explicaremos en el acápite 5.1.1.

5.1.1 Metodología

El corpus SAELE ha sido creado con el propósito de comparar algunas combinaciones de palabras de los aprendices adultos suecos de E/LE (nivel situado entre el A2-B1 de acuerdo a la escala del MCERL) con las de nativos hispanohablantes y suecos. Específicamente para fijarnos en cómo nuestros aprendices usan las combinaciones de palabras con *que: lo que, de que, algo que, dice que*.

En su elaboración pueden identificarse dos etapas. La primera fue un período preparatorio que permitió la exploración informativa acerca de la elaboración de corpus de aprendices, diseñar las estrategias de recogida del material y elegir instrumentos apropiados para la recopilación de datos durante los dos cursos lectivos y las dos universidades suecas donde se efectuó el estudio. La segunda etapa o período ejecutivo permitió la recolección de datos, su procesamiento, análisis de los resultados, conclusiones y recomendaciones. Ambas etapas exigieron de meticulosidad y control de nuestra parte, así como del entrenamiento en herramientas computacionales específicas y la ayuda de un equipo de trabajo integrado por colaboradores de la Universitat de Barcelona y la universidad de Gotemburgo, cuyos nombres se precisan más adelante.

Las características específicas del corpus SAELE se discuten a continuación siguiendo específicamente las recomendaciones metodológicas de Biber (1993) y Sinclair (2005)⁷⁵.

(1) Contenido del corpus

Textos elaborados por nativos suecos aprendices de español en dos centros suecos de enseñanza superior: la universidad de Gotemburgo y el instituto superior de Jönköping. La base de datos SAELE contiene 135 textos argumentativos escritos por estudiantes universitarios de primer año de E/LE (según el MCERL con nivel que oscila entre el A2-B1). 90 textos fueron producidos por aprendices de la universidad de Gotemburgo y 45 por aprendices del instituto superior de Jönköping. El total de palabras en el corpus es de 37.076.

Los textos constituyen redacciones libres sobre los temas siguientes: ‘Mi peli favorita’, ‘Mi tiempo libre’, ‘El deporte que me gusta practicar’; ‘Vacaciones

⁷⁵ Lozano sigue los diez principios metodológicos de Sinclair para la conformación del corpus CEDEL2 (Lozano, 2008).

inolvidables', 'Recuerdos de Navidad' así como la argumentación de un artículo periodístico. Los temas de las redacciones fueron previamente elegidos por dos profesores del curso de Destrezas integradas que se impartía en la universidad de Gotemburgo durante los cursos lectivos 2008/09 y 2009/10, siendo uno de los mismos la autora de la tesis. Ambos profesores teníamos conocimiento de que las redacciones serían utilizadas para el presente estudio. Los temas de las redacciones estaban relacionados con los que trabajamos en el curso de destreza que cursaban los participantes. Su elección a nuestro juicio posibilita la aparición de estructuras lingüísticas variadas que muestran aspectos de la interlengua del aprendiz y está acorde con la función externa que debe cumplir el texto (Sinclair, 2005: 1). Las redacciones debían ser argumentativas, aunque el nivel en que se encuentran los aprendices y las diferencias individuales entre ellos puede mostrar las limitaciones en el dominio de la lengua escrita de estos.

Los criterios de inclusión/exclusión empleados pretendieron garantizar que los sujetos se encontraran al momento de la recopilación de datos aproximadamente al mismo nivel de la lengua meta, entre el A2-B1. Todos los sujetos estudiaban un curso relativamente similar, aunque en dos universidades diferentes.

(2) Representatividad

Biber (1993: 2,3) y Sinclair (2005: 2) coinciden en que las muestras que sean recopiladas deben ser representativas de la lengua del grupo al cual representan, y por ello fuimos muy cuidadosos en cumplir con este parámetro. En nuestro caso, los textos que recopilamos han sido elaborados por aprendices suecos de E/LE del nivel A2-B1 que cursaban estudios en dos universidades suecas, y este se considera directamente el universo del presente estudio. La totalidad de estudiantes en este nivel, que cursaban estudios equivalentes en las dos universidades suecas era de 177 (universo), en el momento de la recopilación de los datos. La muestra total de participantes fue de 45, lo que representa el 25,4 % de la totalidad del universo y por tanto resulta ser representativa de su universo ($P < 0.005$). Los resultados así alcanzados se consideran estadísticamente extrapolables al universo del estudio.

El diseño adoptado es prospectivo y transversal, es decir, los datos se recopilan en dos muestras equivalentes pero cronológicamente asimétricas (durante dos cursos lectivos diferentes), mientras que se utilizan variables definidas para describir y explicar la relación entre ellas (demostrar la homogeneidad de sus sujetos y por tanto

la equivalencia de las muestras independientemente de que son organizadas en momentos diferentes, por razones prácticas).

(3) Contraste

En el corpus participan aprendices suecos de E/LE de dos universidades diferentes, lo que nos permite contrastar determinadas variables utilizadas en la investigación de acuerdo a nuestros objetivos. En nuestro caso utilizamos el subcorpus de hablantes nativos de español del CEDEL2 y un corpus de nativos suecos, el ARGUS, pero ninguno de ellos forma parte del corpus SAELE.

El contraste en el corpus SAELE se puede constatar en la diversidad de sujetos participantes cuyas competencias en principio son similares, pero no exactamente iguales, aunque los grupos se consideran estadísticamente homogéneos, por lo que pueden verse diferencias entre las características de la IL entre cada uno de los sujetos y entre los sujetos de acuerdo a su universidad de procedencia lo cual se aprecia en el análisis textual que llevamos a cabo en el capítulo 7 (Discusión).

(4) Estructura

La estructura del corpus SAELE pretende garantizar que podamos cumplir los objetivos del presente estudio ya que abarca un nivel relativamente homogéneo de competencia entre sujetos procedentes de dos universidades suecas diferentes. La cantidad de sujetos elegida se ha orientado estrictamente a cumplimentar aspectos de la rigurosidad estadística (representatividad); en el futuro pueden añadirse nuevas muestras equivalentes respetando el mismo diseño, o añadir nuevas con otros niveles de competencia lo que permitiría tener un corpus dividido en diferentes subcorpus, por niveles, de acuerdo al MCREL. Lo más importante es que- de acuerdo con Sinclair, hemos hecho un diseño del corpus que tiene en cuenta la recolección adecuada de los textos y su procesamiento para después poder interpretarlos según nuestros propósitos y validar los resultados (2005: 2,5).

(5) Etiquetado

La anotación del corpus se ha comenzado a realizar con la colaboración de Montse Nofre y Oriol Borrega, especialistas del Servicio de Tecnología Lingüística de la UB, pero la misma se concluirá en una etapa posterior a este estudio por las limitaciones de tiempo. Tomamos la decisión de no hacer el etiquetado como parte del trabajo de la tesis ya que en primer lugar no era necesario para realizar esta investigación, y en

segundo antes de efectuar la anotación es preciso planificar qué es lo que se va a etiquetar a partir de las características de la lengua que se pretendan estudiar (Sinclair, 2005: 5).

En el futuro pretendemos utilizar el programa Eclipse y el procesador AnCoraPipe. Por un lado, la anotación se hará con el programa hsmorfo que tiene tres componentes esenciales: un tokenizador (tokenizer), un lematizador y un etiquetador morfológico. Los ficheros anotados son un texto plano en tres columnas (forma- lema-PoS). (Carmona, J. et al., 1998).

Para la revisión y posterior tratamiento del corpus se utilizará el conjunto de herramientas AnCoraPipe, integradas en la plataforma Eclipse⁷⁶, las que permiten trabajar con ficheros en XML y utilizan el lenguaje Xpath para las búsquedas y los recuentos estadísticos.

En este estudio se ha empleado la versión 5 del procesador de texto WordSmith (Scott, 2010)⁷⁷ por su eficacia en la organización y análisis textual, para confeccionar diferentes listas de palabras del corpus SAELE y de los corpus de control.

Las herramientas específicas de *frecuencia de ocurrencias*, *concordancias*, *colocación*, *plot*, *cluster*, permiten un análisis más amplio, sistemático, profundo, mejor validado y – consecuentemente- aporta una mayor potencialidad de generalizar resultados.

En el procesador textual WordSmith con la herramienta *frecuencia de ocurrencias* se puede confeccionar una lista de palabras en cada uno de los tres corpus con que trabajamos. Esta herramienta nos ofrece además el índice de frecuencia de una determinada palabra (o tipo de palabra *type/token*). La herramienta de *concordancias* posibilita identificar una palabra o combinaciones de palabras en contexto y observar las palabras alrededor de la/las misma (s), identificando sus posiciones acorde con la elección que hagamos. La herramienta *colocación* permite identificar palabras (dos o más) que aparezcan de forma simultánea (co-aparecen) en un texto y determinar la distancia entre los elementos de la colocación que habitualmente varía de cuatro a cinco palabras y la función que ocupan estos elementos lingüísticos en el tipo de lenguaje y texto objeto de análisis. La herramienta denominada *plot* permite ilustrar visualmente la forma de distribución de una palabra en un texto; mientras que las

⁷⁶ Véase en: <http://clic.ub.edu/en/ancorapipi-en> en inglés o <http://clic.ub.edu/es/ancorapipi-es> en español.

⁷⁷ Licencia de la universidad de Gotemburgo.

agrupaciones de tres palabras se observan con auxilio de la herramienta denominada *cluster*.

Otros instrumentos utilizados para realizar los análisis estadísticos han sido el programa PASW Statistic 18⁷⁸ (para verificar datos y validar los resultados), y el calculador estadístico Log-likelihood⁷⁹ para la comparación de índices de frecuencia entre diferentes corpus (Rayson, Berridge & Francis, 2004; Rayson & Garside, 2000).

(6) Autenticidad de la muestra

En el corpus SAELE las muestras de los textos son completas, lo que permite observar y analizar diferentes fenómenos de la IL de los aprendices. Esto es un principio muy significativo ya que si la muestra de la lengua es auténtica esto nos revelará resultados confiables acerca del aspecto que investiguemos (Sinclair, 2005: 7). Los textos fueron enviados por correo electrónico o a través de las plataformas electrónicas denominadas GUL (universidad de Gotemburgo) y Ping Pong (instituto superior de Jönköping).

(7) Documentación

La recogida del material comenzó en el curso lectivo 2008/2009 y concluyó en la primavera de 2010, o sea que se han compilado textos de los sujetos de estudio durante dos cursos lectivos (estudio prospectivo y transversal). Los textos fueron enviados por los sujetos vía correo electrónico o plataforma de las universidades, compilados sucesivamente en una base de datos y analizados en forma estructurada pretendiendo servir de inmediato, pero también en el futuro para observar y analizar determinadas regularidades en la destreza escrita del grupo, por ejemplo, errores más frecuentemente cometidos, de ellos cuáles tienen relación con la LM, LE, etc.

Para la base de datos del corpus SAELE hemos un código de identificación de los textos similar al del corpus ARGUS. Este código está compuesto de GU o JU para la universidad; V o H para el sexo; 1, 2 ó 3 refiriéndose al número del texto por sujeto, una letra que identifica al sujeto y el curso lectivo. Por ejemplo: GUV2R89.

⁷⁸ Este programa y el WordSmith fueron instalados por técnicos del departamento de computación de la universidad de Gotemburgo, a quienes agradecemos su ayuda.

⁷⁹ El calculador estadístico Log-likelihood se utiliza en los estudios basados en LC para comparar índices de frecuencias entre diferentes corpus. Véase en: <http://ucrel.lancs.ac.uk>

Para su almacenamiento los textos han sido codificados en su forma original de envío y archivados, sin datos personales u otros que puedan emplearse para identificar al autor (verificabilidad y anonimato). Se conservaron no obstante los caracteres å, ä y ö para nombres propios de lugares o personas que no tuviesen relación con el autor del texto. Los textos se convirtieron en textos planos sin formato para posibilitar su procesamiento con el programa Eclipse y el procesador AnCoraPipe y WordSmith 5. Se cumple el principio del anonimato y los datos recopilados se utilizarán solamente con fines investigativos.

(8) Equilibrio

En este principio consideramos que primase balance en las muestras recogidas que integran el corpus SAELE (Sinclair, 2005: 9). Quiere decir esto que los textos compilados tienen una extensión similar, los mismos temas y los participantes se encontraban todos en un nivel A2-B1 según el MCERL, de esta forma se garantiza que en el análisis lingüístico que vayamos a realizar estén descartados errores concernientes a la calidad y equilibrio de la muestra. El equilibrio del corpus está en relación con la representatividad que este muestra.

(9) Tema

El tema está en relación con el criterio que utiliza el investigador a la hora de elegir qué fenómeno quiere estudiar y cómo, además de tener claro cuál es el objetivo del estudio y de acuerdo con Sinclair, lo más importante es que sea representativo del registro lingüístico que se estudie (2005: 10). En nuestro caso diseñamos una estrategia de trabajo a partir de los objetivos utilizando el material y método para lograr que el corpus fuese representativo del nivel que suelen tener los aprendices durante este período de estudio (nivel A2-B1 según el MCERL).

(10) Homogeneidad

Este principio nos alerta sobre los llamados textos atípicos en nuestro corpus, y concretamente quiere decir que antes de etiquetar el corpus debemos analizar los textos para eliminar aquellos que no cumplan con nuestro diseño y haya por tanto homogeneidad, como indica Sinclair (2005: 14). En el caso del trabajo con el SAELE para este estudio, no se detectó ningún texto atípico que haya sido necesario eliminar.

Por último, consideramos menester puntualizar que aunque el corpus SAELE sea un minicorpus (37.076 total de palabras) se adecua a los objetivos de este estudio debido

a que en su elaboración hemos tratado de combinar juicio y conveniencia siguiendo a Biber (1993).

5.2 Los corpus de control: ARGUS y CEDEL2

Se han seleccionado dos corpus de control, uno de nativos suecos (subcorpus ARGUS), y otro de nativos hispanohablantes (subcorpus de CEDEL2) con el propósito de observar de forma contrastiva cómo los aprendices suecos de E/LE usan determinadas combinaciones de palabras objeto de estudio en comparación con las equivalentes utilizadas en lengua sueca (por nativos suecos) y las utilizadas por hispanohablantes. La elección de los dos corpus conllevó consigo la búsqueda informativa acerca de los corpus de aprendices en general, y posteriormente la valoración de cuáles serían adecuados para nuestro propósito.

5.2.1 El corpus ARGUS

El corpus ARGUS⁸⁰ fue proporcionado por el profesor Lars Borin del Centro de Tecnología lingüística, denominado *Språkbanken* de la Universidad de Gotemburgo. El mismo fue concebido como un corpus de control en el proyecto *CrossCheck-korpusen*⁸¹, está compuesto de 287 textos argumentativos escritos por alumnos suecos de preuniversitario durante los cursos lectivos 1988/89, 1989/90 y 1999/2000. Los temas que se orientaron en la conformación de sus textos fueron los siguientes: *Stoppa barnarbetet!*/¡Paren el trabajo infantil! (39 textos); *Ärade lyssnare*/Querido oyente (10 textos); *Svordomar och annat fult språk – gift eller krydda?*/¿Lenguaje obsceno y malas palabras, vulgaridad o sabor? (19 textos); *Du är inte ni med mig!*/¡Tú no es lo mismo que usted! (105 textos); *Behövs svenskämnet?*/¿Se necesita el sueco como asignatura? (55 textos); *Hur mycket ska man arbeta egentligen?*/¿Cuánto se necesita trabajar en

⁸⁰ Véase: Argumentera! En presentation av gymnasisters argumenterande texter i databasen ARGUS. (Svenska i utveckling nr 19. FUMS Rapport nr 209).

⁸¹ Véase la información sobre el proyecto "CrossCheck - svensk grammatikkontroll för andraspråksskribenter" en <http://www.csc.kth.se/tcs/projects/xcheck/index.html>

realidad? (3 textos); *Om jag vore arbetsmarknadsminister*/Si yo fuese ministro del trabajo (1 texto) y *För barnboken!*/¡A favor de la literatura infantil! (55 textos).

Todos los textos contienen un código de identificación, como por ejemplo: 99GY0573-MVB3; que indica el año (99); identifica el Instituto preuniversitario de procedencia (GY); el número del texto en el archivo (0573); el sexo (M, en sueco hombre o K mujer) o; la V indica el semestre de primavera (en sueco es *vår*) y finalmente B3 la tarea concreta orientada.

Del total de textos del corpus ARGUS seleccionamos 135 para tener una muestra de control relativamente similar. El corpus de control ARGUS que utilizamos tiene un total de 77.650 palabras.

Consideramos que de acuerdo a su contenido y estructura era el más adecuado para utilizar como corpus de control de lengua sueca.

5.2.2 El corpus CEDEL2

El subcorpus de nativos hispanohablantes (de control a los efectos del presente estudio) forma parte del CEDEL2 (Corpus Escrito del Español L2) que es un proyecto de investigación en colaboración entre la Universidad Autónoma de Madrid y la Universidad de Granada que investiga aspectos de la lengua escrita durante el aprendizaje de E/LE en sujetos que tienen el inglés como lengua materna, L1 inglés-L2 español, de todos los niveles: principiante, intermedio, avanzado⁸².

El material del corpus se recopila online a través de formularios electrónicos. Contiene en la actualidad un total de 191.331 palabras en los textos producidos por un total de 660 sujetos anónimos⁸³, se respeta la confidencialidad y su uso es con fines investigativos.

Los sujetos rellenaron previamente un formulario proporcionando información acerca de sus datos personales (edad, sexo, variedad del español que habla según su país, lengua materna, lengua materna del padre, de la madre y la que se habla en casa,

⁸² Su utilización a los efectos del presente estudio ha sido una cortesía de su autor Cristobal Lozano de la Universidad de Granada. Agradecemos también a Mar Cruz Piñol de la Universitat de Barcelona por habernos hecho la sugerencia acerca del subcorpus de aprendices hispanohablantes del CEDEL2.

⁸³ Véase la información técnica actual sobre el CEDEL2 en: <http://www.uam.es/proyectosinv/woslac/collaborating.htm>

dominio de otras lenguas extranjeras (hablar, comprender, leer y/o escribir clasificándolas en niveles avanzado, intermedio o principiante).

A continuación los participantes eligen un tema de redacción entre 12 posibles: ‘¿Cómo es la región dónde vives?’; ‘Habla de una persona famosa’; ‘Resume una película que has visto recientemente’; ‘¿Qué hiciste el año pasado durante las vacaciones?’; ‘¿Cuáles son tus planes para el futuro?’; ‘Describe un viaje que has hecho recientemente’; ‘Cuenta una experiencia que hayas vivido’; ‘Habla del problema del terrorismo en el mundo’; ‘¿Qué opinas de la nueva ley antitabaco?’; ‘¿Crees que las parejas gay tienen derecho a casarse y adoptar niños?’; ‘¿Crees que la marihuana se debería legalizar?’ y ‘Analiza los principales aspectos de la inmigración’.

Al escribir la composición debían especificar si se habían documentado sobre el tema con anterioridad y si la respuesta era afirmativa precisar si había sido a través de internet, periódico o revista, libros y artículos, programas de televisión o radio u otros. Aunque no se limitaba el número de palabras, se recomendaba escribir al menos varios párrafos y evitar el uso de diccionarios y/o correctores ortográficos⁸⁴.

De este corpus también seleccionamos 135 textos para tener una muestra de control equivalente al corpus SAELE. El subcorpus de CEDEL2 que utilizamos tiene un total de 39.237 palabras. Los textos seleccionados son de la parte del corpus CEDEL2 que también es de control, lo que quiere decir que los textos fueron elaborados por sujetos cuya lengua materna es el español.

Resumiendo, los dos corpus elegidos como de control en nuestro estudio son apropiados para realizar el estudio contrastivo de las combinaciones de palabras con *que*: *lo que*, *de que*, *algo que*, *dice que* por su contenido, representatividad, tamaño, estructura, equilibrio, documentación y homogeneidad.

5.3 La encuesta y sus parámetros metodológicos

La encuesta (Anexo 5) fue elaborada basándonos en las recomendaciones de Creswell (Creswell, 2003), modificadas según aspectos considerados en investigaciones anteriores vigentes para corpus de aprendices como por ejemplo: el

⁸⁴ Véanse los formularios electrónicos en: <http://www.uam.es/proyectosinv/woslac/start-spanish.htm>

ICLE, *International Corpus Of Learner English* (Granger et al. 2002) el corpus ARGUS y el corpus CEDEL2⁸⁵ (Lozano, 2008).

El objetivo del diseño fue aplicar categorías que permitieran en primer lugar poder recoger los datos sociolingüísticos de los informantes para con posterioridad describir el perfil del aprendiz sueco de E/LE que se encuentra en un nivel A2-B1 acorde con el MCERL que participa en este estudio (véase los objetivos de la tesis en la introducción).

Las categorías utilizadas fueron:

- (a) Edad: esta categoría nos ofrece información sobre la edad promedio de los participantes de las dos universidades suecas que forman parte del estudio (véase el diagrama 5).
- (b) Sexo: nos indica la tendencia existente respecto a mujeres y hombres que estudian E/LE en el nivel superior (véase el diagrama 5.)
- (c) Lengua materna: un importante indicador ya que de acuerdo a nuestro estudio tenemos una muestra de participantes cuya lengua materna es solamente el sueco.
- (d) Tiempo de estudio de E/LE: Aquí se clasifica el sujeto con estudios entre menos de dos años y seis años y podemos conocer cuánto tiempo el aprendiz ha estado estudiando la LE en un contexto educacional.
- (d) Lugar donde había cursado estudios de E/LE: en Suecia, otro país no hispanohablante, España, Latinoamérica). Esta categoría nos sirve para conocer aún más sobre qué conocimientos reales de E/LE tienen los participantes y apreciar en qué medida ellos han estado en contacto con la lengua española en un contexto donde esta es la L1 (véase el diagrama 6).
- (e) Eventual tiempo de estudio de E/LE fuera de Suecia: lo desglosamos entre menos de dos años a seis años y los resultados nos ofrecen más elementos que pueden explicar en alguna medida el nivel de competencia en la LE del aprendiz.
- (f) Intensidad de estudio semanal de E/LE: la desglosamos entre menos de 5 horas hasta más de veinte horas a la semana. Esta categoría se desprende de la anterior al poder ganar más elementos sobre la intensidad de estudios realizados en la LE.

(g) Estudio de otra (s) LE (específicamente inglés, francés, alemán, italiano, portugués u 'otras'). Nos permite conocer cuántas lenguas han estudiado los aprendices y a partir de ello deducir si poseen o no determinadas estrategias en el aprendizaje de LE, además de posiblemente explicar algunas interferencias cuando escriben los textos en E/LE.

(h) Tiempo de estudio de otra (s) LE (desglosado entre menos de dos hasta seis años): esta categoría nos permite descubrir qué otras lenguas han estudiado los participantes para también poder analizar si alguna de ellas influye o no en el aprendizaje de la nueva LE (véase el diagrama 7).

(i) Autovaloración de las destrezas escritas en E/LE: específicamente comprensión de lectura (excelente, muy bien, bien, regular o mal) y autovaloración de la expresión escrita (excelente, muy bien, bien, regular o mal): Esta última categoría de la encuesta tiene que ver con la autoconfianza del aprendiz que puede coincidir o no con la evaluación que realizamos de los textos producidos por ellos (véase el diagrama 8)

Otro elemento relevante a destacar por su implicación ética es el referido a la distribución de la encuesta y sus permisos, por lo que explicamos a continuación cómo realizamos esta tarea.

En la primera parte del estudio la encuesta fue distribuida en el aula durante los cursos lectivos 2008/09 y 2009/10 a todos los estudiantes que voluntariamente decidieron participar, y que estaban matriculados en alguno de los cursos elegidos. Esta información se dio (y la encuesta se distribuyó) en la universidad de Gotemburgo en los semestres de primavera y otoño (se imparte un curso en cada semestre). En el instituto superior de Jönköping solamente en el semestre de otoño (un solo curso durante todo el curso lectivo).

Los sujetos que forman parte de este estudio dieron su consentimiento escrito de participación, después de haber sido previamente informados sobre los objetivos de la tesis, datos personales de los tutores y la autora (Anexo 3) y haber recibido garantías de anonimato. En el Anexo 4 se presenta el certificado de consentimiento informado de participación voluntaria, copia del cual se entregó a cada estudiante. Los originales han sido archivados para eventual verificación.

5.3.1 Los participantes en el estudio

A los efectos de la encuesta fueron los criterios de inclusión: estudiantes que cursaban estudios de E/LE y estaban matriculadas en el curso *Grundkurs I Spanish/Introductory Course I* en la universidad de Gotemburgo⁸⁶ durante el curso lectivo 2008/09 (62 sujetos: 23 en el curso del otoño de 2008 y 39 en el curso de la primavera de 2009), o en el curso lectivo 2009/10 (64 sujetos: 35 en el otoño y 29 en la primavera). O bien matriculados en el curso *Spanska 1-30 hp* del curso lectivo 2008/09 (27 sujetos) o del curso lectivo 2009/10 (24 sujetos) en la facultad de educación y comunicación del instituto superior de Jönköping⁸⁷. La participación fue voluntaria.

En total había 89 matriculados en ambas universidades en 2008/2009 y 88 en el 2009/2010. De ellos accedieron a participar en la encuesta 27 sujetos en 2008/2009 (30% del universo) y 27 sujetos en el curso 2009/2010 (31 % del universo). En nuestro estudio consideramos como universo, el total de sujetos matriculados en los cursos lectivos/año que hemos hecho referencia.

De los 27 sujetos participantes en la encuesta aplicada en el curso 2008/2009 21 (13 en Gotemburgo y 8 en Jönköping) fueron incluidos para la conformación del corpus SAELE. Otras lenguas maternas (no sueco) representadas en la muestra original fue la razón del 50% de pérdida estadística, el resto fue debido a participantes que se registraron como interesados pero no entregaron la encuesta.

En el cuadro 5 se recopilan datos estadísticos sobre los participantes en la encuesta. La edad promedio de los que tenían el sueco como LM osciló entre 19-63 años. En Gotemburgo predominó una muestra de personas más jóvenes y homogéneamente distribuida en sexos, mientras que el 90 % de la muestra en Jönköping estuvo representada por mujeres. Esta peculiaridad entre las muestras indica probablemente

⁸⁶ Información obtenida de Álvaro Foresti, jefe de estudios de la sección de español. Departamento de lengua y literaturas, Facultad de Humanidades. Universidad de Gotemburgo.

⁸⁷ Información obtenida de Kerstin Wigertz, orientadora de estudios. Facultad de educación y comunicación. Instituto superior de Jönköping.

MATERIAL DE TRABAJO

diferencias demográficas en la población asentada en ciudades grandes y medianas en Suecia⁸⁸.

Cuadro 5. Participantes curso lectivo 2008/09

Curso lectivo 2008/09	Universidad de Gotemburgo	Instituto superior de Jönköping
Participantes	Total=18 (Pérdida=5)	Total=9 (Pérdida=1)
Sujetos descartados (Pérdida)	Total=5 2 lengua materna portugués 3 no entregaron la encuesta	Total=1 Lengua materna farsi
Edad promedio de los seleccionados	33 años	47 años
Sexo	Femenino=7 Masculino=6	Femenino=7 Masculino=1
Lengua materna	Sueco	Sueco

El cuadro 6 muestra datos correspondientes a los 27 sujetos del curso lectivo 2009/10 que participaron en la encuesta (17 de Gotemburgo y 10 de Jönköping), la pérdida estadística en este caso estuvo representada por 3 sujetos que no tenían el sueco como LM. De los que integraron el SAELE, la edad fue de 19-63 años, con un promedio de 21 años en la universidad de Gotemburgo y de 33 en la Facultad de educación y comunicación del instituto superior de Jönköping. En la universidad de Gotemburgo se registró nuevamente la misma cantidad de participantes entre

⁸⁸ Véase el informe de la Dirección nacional de educación superior sueca:
<http://www.hsv.se/download/18.ac0b56a12a53b8af8580001119/1010R-universitet-hogskolor-arsrapport-2010-utbild-grund-avancerad.pdf> [03-03-2011].

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

mujeres (8) y hombres (8), mientras que en la Facultad de educación y comunicación del instituto superior de Jönköping el 60% de los sujetos fueron mujeres.

Cuadro 6. Participantes curso lectivo 2009/10

Curso lectivo 2009/10	Universidad de Gotemburgo	Instituto superior de Jönköping
Participantes	Total=17 (Pérdida=1)	Total=10 (Pérdida=2)
Sujetos descartados (Pérdidas)	Total=1 (lengua materna croata)	Total=2 (lengua materna español, 1 no entregó la encuesta)
Edad promedio de los seleccionados	22 años	34 años
Sexo	Femenino=8 Masculino=8	Femenino= 5 Masculino=3

Los Cuadros 7 y 8 describen el tiempo, lugar, duración e intensidad de estudios previos de E/LE entre los sujetos cuyos textos integraron el SAELE, dividido en dos subgrupos. Se evidencia que el tiempo previo de estudio de E/LE en los sujetos procedentes de ambos subgrupos y ciudades fue de aproximadamente 3 años. El 70% de los sujetos de la universidad de Gotemburgo en el curso lectivo 2008/09 respectivamente el 50% en el instituto superior de Jönköping, así como el 25% respectivamente el 50% de los correspondientes al curso 2009/10 en el mismo orden hicieron esos estudios en un país hispanohablante.

MATERIAL DE TRABAJO

Cuadro 7. Relación entre tiempo de estudio de E/LE, lugar, duración (país hispanohablante) e intensidad

Curso lectivo 2008/09	Universidad de Gotemburgo	Instituto superior de Jönköping
Tiempo de estudio de E/LE	6 sujetos: de 2-4 años 4 sujetos: < de dos años 2 sujetos: de 5-6 años 1 sujeto: > de 6 años	4 sujetos: de 2-4 años 3 sujetos: < de 2 años 1 sujeto: > de seis años
Lugar de estudios	7 sujetos: Suecia 4 sujetos: España 5 sujetos: Hispanoamérica	5 sujetos: Suecia 4 sujetos: España 1sujeto Hispanoamérica
Duración estudios país hispanohablante	7 sujetos: < de dos años 1 sujeto: 2-4 años (85 % menos de 2 años)	4 sujetos: < de dos años 1 sujeto: 2-4 años (75% menos de 2 años)
Intensidad de estudios	3 sujetos= 20 horas/semana o más 1 sujeto= 10-20 horas/semana 2 sujetos= 5-9 horas/semana	3 sujetos= 20 horas/semana o más 2 sujetos= 10- 20 horas/ semana

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

Cuadro 8. Relación entre tiempo de estudio de E/LE, lugar, duración (país hispanohablante) e intensidad

Curso lectivo 2009/10	Universidad de Gotemburgo	Instituto superior de Jönköping
Tiempo de estudio de E/LE	4 sujetos: 2-4 años 8 sujetos: < de dos años 2 sujetos: de 5-6 años 2 sujetos: > de 6 años	3 sujetos: 2-4 años 2 sujetos: < de 2 años 2 sujetos: 5-6 años 1 sujeto: > de seis años
Lugar de estudios	12 sujetos: Suecia 3 sujetos: España 1sujeto: Hispanoamérica	5 sujetos: Suecia 2 sujetos: España 3 sujetos: Hispanoamérica
Duración de estudios en un país hispanohablante	5 sujetos: < de dos años	3 sujetos:< 2 años
Intensidad de estudios	1 sujeto: 20 horas/semana o más 3 sujetos:10-20 horas/semana 1 sujeto: < 5 horas/semana	1 sujeto: 10- 20 horas/ semana 2 sujetos:<5horas/se- mana

Debemos señalar que aunque en este capítulo nos referimos a las categorías de la encuesta elaborada e incluso mostramos algunos de sus resultados, estos los debatiremos en el capítulo de discusión (7) con la ayuda de diagramas confeccionados con este propósito debido a que partimos que el sujeto participante autor de los textos que forman parte del corpus SAELE como ser social está influido

MATERIAL DE TRABAJO

por factores sociolingüísticos que en buena medida determinan el proceso individual de aprendizaje y sus resultados.

CAPÍTULO 6. RESULTADOS CUANTITATIVOS DE LOS CORPUS

El presente capítulo presenta los resultados cuantitativos del Corpus SAELE (6.1) y los registrados en los corpus de control ARGUS (subcorpus de nativos suecos) (6.2), CEDEL2 (subcorpus de nativos hispanohablantes) (6.3) y la comparación de frecuencia entre corpus (6.4) según las herramientas del procesador de texto WordSmith y el calculador estadístico Log-likelihood. Los cuadros, figuras y diagramas pretenden facilitar la exposición de los elementos más relevantes de los resultados.

6.1 Resultados del corpus SAELE

El corpus SAELE está compuesto de 135 textos, redactados por 27 mujeres y 18 hombres, aprendices suecos de primer año universitario de E/LE (en nivel aproximado entre el A2 y B1 según el MCERL). De los 135 textos, 90 fueron recopilados en la universidad de Gotemburgo y 45 en el instituto Superior de Jönköping. En el presente trabajo se analizan el tipo de uso y frecuencia en el corpus de estudio, comparando con los subcorpus de control ARGUS (nativos suecos) y CEDEL2 (nativos hispanohablantes).

El corpus SAELE tiene un total de 37.076 palabras. Las palabras más frecuentemente utilizadas son: *de* (con frecuencia de aparición de 1508 = 4.06%), *la* conjunción *y* (con frecuencia de 1299 = 3.50%), *la* (frecuencia de 1211 = 3,26%) *en* (con frecuencia de 1139 = 3,07%), y *que* (con una frecuencia de 1094 = 2.95%).

En la figura 1 se muestra la lista de palabras del corpus SAELE (líneas de la 1-35) el número que ocupan en la lista, su frecuencia de aparición, y porcentaje así como el número de textos en que aparecen.

RESULTADOS CUANTITATIVOS DE LOS CORPUS

N	Word	Freq.	%	Texts
1	DE	1.508	4,07	135
2	Y	1.299	3,50	134
3	LA	1.211	3,27	134
4	EN	1.139	3,07	135
5	QUE	1.094	2,95	134
6	EL	970	2,62	131
7	LOS	650	1,75	130
8	UN	623	1,68	132
9	ES	619	1,67	125
10	A	613	1,65	128
11	UNA	483	1,30	124
12	NO	372	1,00	114
13	CON	365	0,98	122
14	POR	356	0,96	111
15	#	325	0,88	92
16	PARA	321	0,87	120
17	SE	307	0,83	94
18	LAS	295	0,80	109
19	MUY	295	0,80	102
20	ME	291	0,78	91
21	PERO	278	0,75	108
22	MI	233	0,63	74
23	MÁS	229	0,62	95
24	DEL	222	0,60	82
25	CUANDO	205	0,55	91
26	COMO	200	0,54	92
27	TAMBIÉN	182	0,49	92
28	SU	176	0,47	59
29	ERA	173	0,47	60
30	MUCHO	152	0,41	73
31	LO	141	0,38	74
32	AL	140	0,38	72
33	HAY	136	0,37	76
34	PORQUE	131	0,35	72
35	SON	128	0,35	66

Figura 1. Lista de palabras del corpus SAELE

En la figura 2 se presentan las 10 primeras líneas de concordancias con *que*, de un total de 1094 contenidas en el corpus SAELE. El análisis de las líneas de concordancias permiten establecer la relación en contexto de (en este caso) *que* con las palabras que la preceden o siguen. En estas primeras 10 líneas se observa que *algo* precede a *que* en las líneas 3 y 6 formando la combinación *algo que*. Además a la derecha y seguidamente de *que* podemos ver verbos como por ejemplo: trabajar (línea 1), estar (línea 7) y querer (línea 9).

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

N	Concordance	Set Tag	Word #	t. #	os. a. #	os. d. #	os. t. #	os. l. #	File	%	
1	hombre, Edward Lewis (Richard Gere) <i>que</i> trabaja como un hombre de		32	2	2%	0	9%	0	9%	GUH1A89.txt	9%
2	rico y afortunado. Pero hay una cosa <i>que</i> falta en su vida, el amor. Edward va		47	3	0%	0	3%	0	3%	GUH1A89.txt	13%
3	Roberts) para ser su prometida. Algo <i>que</i> Edward no cuenta con es que el va		87	5	3%	0	4%	0	4%	GUH1A89.txt	24%
4	Algo que Edward no cuenta con es <i>que</i> el va a amorarse con ella, una mujer		93	5	8%	0	6%	0	6%	GUH1A89.txt	26%
5	Pretty Woman es una película <i>que</i> se trata muchos temas diferentes y		114	6	4%	0	1%	0	1%	GUH1A89.txt	32%
6	son muy diferentes pero igual hay algo <i>que</i> los mantiene unido; la soledad y el		198	10	8%	0	5%	0	5%	GUH1A89.txt	54%
7	famosas y compra ropa lujosa. Yo creo <i>que</i> esta oportunidad, a hacer compras		309	16	7%	0	5%	0	5%	GUH1A89.txt	85%
8	de todas las chicas. Es una película <i>que</i> se dirige a espectadores jóvenes,		332	17	5%	0	2%	0	2%	GUH1A89.txt	92%
9	primero a chicas pero también a gente <i>que</i> quiere soñarse a un mundo donde		345	17	3%	0	5%	0	5%	GUH1A89.txt	96%
10	pero me gusta viajar y hay varios sitios <i>que</i> me gustaría visitar. Pues bien,		23	0	8%	0	9%	0	9%	GUH1AA910.txt	8%

Figura 2. Líneas de concordancias con *que*. Corpus SAELE

Si separamos el corpus SAELE en dos minicorpus, el SAELEgu (universidad de Gotemburgo) y el SAELEju (instituto superior de Jönköping) reparamos que en el primero aparece *que* ocupando el lugar quinto en la lista de palabras mientras que en SAELEju ocupa el cuarto lugar (cuadro 11). La frecuencia de aparición de *que* en el minicorpus SAELEgu es de 671 (2,81 %), aparece en el total de los 90 textos allí recopilados; mientras que aparece en una frecuencia de 423 (3,21 %) y aparece en 44 del total de 45 textos en el minicorpus SAELEju. El minicorpus SAELEgu tiene un total de palabras de 23.881 y el SAELEju de 13.195.

Cuadro 9. *Que* en la lista de palabras. SAELEju y SAELEgu

N.	Palabra	Frecuencia	%	Textos
4(SAELEju)	Que	423	3,21	44
5(SAELEgu)	Que	671	2,81	90

Al realizar una búsqueda de la presencia de las combinaciones de palabras con *que*: *lo que*, *de que*, *algo que* y *dice que* (Cuadro 10) constatamos que las concordancias con *lo que* son las más frecuentes y con *algo que* las menos frecuentes en el corpus SAELE.

Cuadro 10. Concordancias de combinaciones de palabras con *que*. Corpus SAELE

Concordancias	<i>Lo que</i>	<i>De que</i>	<i>Algo que</i>	<i>Dice que</i>
Total	40	17	13	15

RESULTADOS CUANTITATIVOS DE LOS CORPUS

En la figura 3 se presentan 15 de las 40 líneas de concordancia con *lo que* del corpus SAELE. Se destaca en la línea de concordancia 8 la expresión idiomática: cueste lo que cueste y *lo que* usado como conector textual en las líneas 1, 11 y 12.

N Concordance

- 1 todos los días están esperando afuera. **Lo que** necesitan es un plato de comida
- 2 hay 34 millones de hispanohablantes, **lo que** representa casi el 20 % de la
- 3 usan habitualmente el Internet **lo que** representa algo más del 6% de
- 4 ha gustado Borás y uno nunca sabe **lo que** va a pasar en el futuro. A lo mejor
- 5 y ver el cómic del algunas situaciones **lo que** hace esta película tan realista y
- 6 piensa que eso no es el caso sino, **lo que** les pone en desacuerdo es el
- 7 precio. Chávez está dispuesto a pagar **lo que** la empresa invirtió inicialmente
- 8 me decidí que voy a obtenerla, cueste **lo que** cueste! Parecia como tenia mala
- 9 día. Es un recuerdo que nunca olvidaré **lo que** experimenté allí y también y
- 10 y ver el cómic del algunas situaciones **lo que** hace esta película tan realista y
- 11 especiales y escenas de acción. **Lo que** me gusta mucho en las primeras
- 12 chistosas, desaparecen poco a poco. **Lo que** nos queda son las escenas de
- 13 también están esperando y escuchan **lo que** dice Forrest Gump de su vida.
- 14 va en una excursión con los chicos, **lo que** fue la razón de su despido
- 15 pero no puedo escribir sobre todo **lo que** pasó durante mi tiempo allí y por

Figura 3. 15 de las 40 líneas de concordancias con *lo que*. Corpus SAELE

La figura 4 muestra las 17 líneas de concordancias con *de que* del corpus SAELE. Se destacan las expresiones: a pesar de que (líneas 2,4, 12 y 13) y el hecho de que (línea 5).

N Concordance

- 1 de crímenes. Está presentada una cifra **de que** unas 350.000 mujeres en España
- 2 donde están Guido y Giosué, a pesar **de que** ella no era judía. Desde aquí la
- 3 esperanza para la vida. La esperanza **de que** hay algo más poderoso que la
- 4 eran muy humildes y amables, a pesar **de que** han experimentado cosas
- 5 era maravilloso, ni siquiera el hecho **de que** yo casi no había dormido nada
- 6 – pero no más. También es un signo **de que** los tiempos cambian.
- 7 de la terapia de grupo. En caso **de que** se diga algo negativo es quizás
- 8 pero creo que nos pueden hacer mejor **de que** han hecho hasta que ahora.
- 9 del sur. Viajé con un amigo y, a pesé **de que** no hemos proyectado algunos
- 10 puedo sentir la velocidad y el sentido **de que** se volar. Otra cosa curiosa es
- 11 de Fiona. Pero en el artículo se habla **de que** "las agentes femininas imponen
- 12 el primer puesto en la lista a pesar **de que** el nivel de pobreza y desigualdad
- 13 hace todo para ayuda su hijo. A pesar **de que** ella tiene un tiempo duro,
- 14 Además, Katarína Munther se pregunta **de que** importancia su raza tiene en ese
- 15 retiran su dinero por miedo **de que** la situación económica se
- 16 se tratan lo que fue mal y caso **de que** alguien se chiva a la policía. El
- 17 lo comprendieron. Cinco horas antes **de que** tuve que ir al aeropuerto me

Figura 4. 17 (totalidad) líneas de concordancias con *de que*. Corpus SAELE

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

Las concordancias con *algo que* aparecen en 13 líneas (Figura 5); mientras que *dice que* (Figura 6) apareció en 15 líneas, en el Corpus SAELE.

N Concordance

- 1 eso es un poquito cliché-montado, pero *algo que* me gusta. Nosotros gustamos
- 2 (Julia Roberts) para ser su prometida. *Algo que* Edward no cuenta con es que
- 3 dos son muy diferentes pero igual hay *algo que* los mantiene unido; la soledad
- 4 Después, el episodio continua con *algo que* ocurre a los amigos y como se
- 5 bueno y ya tiene que concentrarse, *algo que* no es tan común que yo quiero.
- 6 diferentes pescados "semi-exóticos", *algo que* me gustaba. Suecia, en
- 7 hace relajado. Escucha la música es *algo que* me gusta mucho y ir en
- 8 los miembros de OCDE está subiendo, *algo que* puede indicar que dentro de
- 9 contaminado por recursos frescos, *algo que* se aprobó el Congreso del país.
- 10 Europa y el Banco Central Económico, *algo que* puede ayudar a crear un
- 11 hace relajado. Escucha la música es *algo que* me gusta mucho y ir en
- 12 (Julia Roberts) para ser su prometida. *Algo que* Edward no cuenta con es que
- 13 dos son muy diferentes pero igual hay *algo que* los mantiene unido; la soledad

Figura 5. Líneas de concordancias con *algo que*. Corpus SAELE

N Concordance

- 1 del gobierno llamada Teresa Vega *dice que* la financiación pública de los
- 2 de los albergues no va seguir así. *Dice que* las empresas tienen que tomar
- 3 de la organización de Jaén Acoga *dice que* no se debe de cerrar los
- 4 en España, María José Barahona. Ella *dice que* eso ocurre como un efecto de
- 5 de las que el 70% son extranjeras. Se *dice que* las rutas utilizadas por las
- 6 EE.UU., del Reino Unido y de Holanda, *dice que* Venezuela está obligado no
- 7 vegetales. El doctor Lennart Mucke *dice que* necesitan muchas más
- 8 El Ministerio de Industria *dice que* el sector ha registrado su peor
- 9 del gobierno en los países en Europa. *Dice que* Alemania es el país europeo
- 10 son embarazadas etc. Una tabla *dice que* en Europa, España es un de
- 11 pero es una palabra relativa. Si se *dice que* es una libertad cuando se
- 12 de los hoteles y el ocupación. *Dice que* los turistas han pagado esto
- 13 de los apartamentos son de Inglaterra. *Dice que* el destino más preferido en
- 14 medio eran las cifras 6,1% y 4,5% . *Dice que* España era el centro del
- 15 20 años en el prisión y él como Andy *dice que* es inocente. "Red", bueno y de

Figura 6. Líneas de concordancias con *dice que*. Corpus SAELE

La aplicación de la herramienta concord en el corpus SAELE permitió detectar un total de 290 colocaciones con *que*. La figura 7 muestra la lista de colocaciones con *que* (líneas 1-42). Entre ellas destacan en la línea 18 la colocación *lo que*; en la línea 2 *de que*; en la 30 *algo que*; y en la 42 *dice que*. La figura muestra también el número de textos donde las colocaciones aparecen, por ejemplo en la línea 2 *de que* (se refleja que la combinación aparece en el Corpus en 103 de los 135 textos que lo componen),

RESULTADOS CUANTITATIVOS DE LOS CORPUS

mientras que *dice que* (línea 42) apareció en 21 de los 135 en total que tiene el corpus SAELE.

N	Word	With	elation	Texts	Total	total	Left	al	Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	QUE	que	0,000	134	1.232	69	69	20	19	17	17	13	0	1.094	0	13	17	19	20	
2	DE	que	0,000	103	352	192	160	31	45	54	45	17	0	1	30	44	47	38		
3	LA	que	0,000	96	292	146	146	30	28	36	43	9	0	43	23	30	28	22		
4	ES	que	0,000	97	266	114	152	16	19	26	29	24	0	60	22	24	24	22		
5	Y	que	0,000	102	261	164	97	27	29	46	40	22	0	0	4	29	26	38		
6	EL	que	0,000	98	253	110	143	25	28	24	31	2	0	44	28	29	19	23		
7	EN	que	0,000	104	250	92	158	17	25	33	14	3	0	15	47	32	34	30		
8	LOS	que	0,000	79	186	96	90	20	8	23	37	8	0	36	14	12	14	14		
9	UN	que	0,000	90	179	97	82	13	20	26	38	0	0	8	27	22	11	14		
10	A	que	0,000	82	153	62	91	12	15	21	11	3	0	7	27	30	15	12		
11	UNA	que	0,000	75	141	79	62	12	10	15	42	0	0	2	25	15	8	12		
12	SE	que	0,000	67	140	38	102	10	6	15	7	0	0	74	10	7	7	4		
13	NO	que	0,000	59	98	30	68	5	10	5	10	0	0	31	11	9	10	7		
14	POR	que	0,000	53	93	49	44	11	9	12	6	11	0	1	5	16	12	10		
15	MÁS	que	0,000	56	89	39	50	3	5	2	23	6	0	4	10	13	9	14		
16	LAS	que	0,000	46	89	34	55	7	4	3	18	2	0	18	7	11	10	9		
17	ME	que	0,000	52	87	33	54	7	5	4	17	0	0	41	3	3	3	4		
18	LO	que	0,000	51	80	57	23	4	5	2	6	40	0	9	5	4	1	4		
19	CON	que	0,000	54	76	38	38	6	11	14	5	2	0	0	9	11	6	12		
20	MUY	que	0,000	46	63	19	44	4	8	1	6	0	0	12	12	12	8			
21	PARA	que	0,000	51	60	28	32	7	7	4	4	6	0	4	2	3	15	8		
22	PERO	que	0,000	44	57	37	20	5	9	10	10	3	0	0	1	3	12	4		
23	HAY	que	0,000	38	55	39	16	11	3	15	10	0	0	11	1	1	0	3		
24	ERA	que	0,000	28	52	9	43	0	2	5	0	2	0	16	11	10	5	1		
25	SU	que	0,000	33	51	15	36	3	3	3	6	0	0	6	2	8	9	11		
26	TIENE	que	0,000	35	50	21	29	2	3	1	1	14	0	11	8	4	4	2		
27	TAMBIÉN	que	0,000	39	50	34	16	5	6	10	7	6	0	8	3	0	2	3		
28	DEL	que	0,000	33	49	27	22	6	7	4	10	0	0	0	0	5	11	6		
29	MI	que	0,000	33	47	21	26	3	3	10	4	1	0	1	8	5	1	11		
30	ALGO	que	0,000	24	41	24	17	1	3	4	3	13	0	0	9	3	1	4		
31	CREO	que	0,000	31	41	39	2	1	0	0	0	38	0	2	0	0	0	0		
32	ESTÁ	que	0,000	24	40	7	33	3	1	3	0	0	0	15	5	4	2	7		
33	MUCHO	que	0,000	32	39	10	29	4	0	3	1	2	0	0	6	13	5	5		
34	SON	que	0,000	25	36	14	22	3	0	6	3	2	0	4	3	3	8	4		
35	LLAMA	que	0,000	20	32	3	29	1	0	0	2	0	0	0	28	1	0	0		
36	PELÍCULA	que	0,000	14	31	16	15	3	2	0	2	9	0	0	3	7	3	2		
37	YO	que	0,000	20	31	13	18	1	3	0	9	0	0	14	2	0	0	2		
38	HA	que	0,000	21	31	10	21	5	1	0	4	0	0	8	2	8	2	1		
39	GUSTA	que	0,000	21	30	7	23	1	0	4	1	1	0	0	21	2	0	0		
40	AL	que	0,000	26	30	14	16	2	6	3	3	0	0	2	6	2	3	3		
41	PUEDE	que	0,000	17	28	12	16	4	1	2	5	0	0	4	3	6	3	0		
42	DICE	que	0,000	21	28	17	11	1	1	0	0	15	0	6	3	0	0	2		

Figura 7. Colocaciones con *que* (de la 1-42). Corpus SAELE

La figura 8 compara la lista de palabras del corpus SAELE con el CEDEL2 (corpus de control de nativos hispanohablantes). La palabra *que* (línea 39) aparece en el SAELE con una frecuencia de 1094 (2.94%) y de 1550 (3.95%) en el CEDEL2.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

N	Key word	Freq.	%	Freq.	RC.	%	eyness	P	emmas	Set
1	#	325	0,88	110	0,28		124,29	000000		
2	GUSTA	120	0,32	14	0,04		102,30	000000		
3	TAMBIÉN	182	0,49	54	0,14		80,98	000000		
4	MUY	295	0,80	128	0,33		77,98	000000		
5	SUECIA	62	0,17	2			74,42	000000		
6	CUANDO	205	0,55	85	0,22		58,43	000000		
7	DIVERTIDO	39	0,11	0			56,33	000000		
8	ERA	173	0,47	67	0,17		54,84	000000		
9	MUCHO	152	0,41	56	0,14		51,77	000000		
10	CARTAGENA	26	0,07	0			37,55	000000		
11	ESPAÑA	99	0,27	34	0,09		37,02	000000		
12	FORREST	24	0,06	0			34,66	000010		
13	ES	619	1,67	461	1,17		33,49	000042		
14	EJEMPLO	44	0,12	7	0,02		32,06	000120		
15	MÁLAGA	21	0,06	0			30,32	000336		
16	DICE	40	0,11	6	0,02		30,13	000376		
17	LLAMA	41	0,11	7	0,02		28,64	000841		
18	MUJERES	51	0,14	12	0,03		28,27	001027		
19	FUIMOS	49	0,13	11	0,03		28,23	001048		
20	TIENE	100	0,27	42	0,11		27,85	001283		
21	PAÍSES	49	0,13	12	0,03		26,24	002989		
22	PERO	278	0,75	182	0,46		26,15	003134		
23	NAVIDAD	18	0,05	0			25,99	003399		
24	LIBRE	38	0,10	7	0,02		25,29	004906		
25	ARTÍCULO	22	0,06	1			24,87	006097		
26	POBREZA	17	0,05	0			24,55	007220		
27	DESIGUALDAD	17	0,05	0			24,55	007220		
28	ESQUIAR	17	0,05	0			24,55	007220		
29	TIEMPO	111	0,30	53	0,14		24,43	007670		
30	MAS	35	0,09	6	0,02		24,39	007845		
31	LOS	650	1,75	517	1,32		24,03	009463		
32	SE	307	0,83	464	1,18	-24,16	008830			
33	TE	3		31	0,08	-25,29	004901			
34	AL	140	0,38	252	0,64	-26,55	002533			
35	O	85	0,23	176	0,45	-27,55	001501			
36	LO	141	0,38	257	0,65	-28,20	001066			
37	NI	4	0,01	44	0,11	-36,80	000000			
38	A	613	1,65	896	2,28	-39,34	000000			
39	QUE	1.094	2,95	1.550	3,95	-57,30	000000			
40	DE	1.508	4,07	2.135	5,44	-79,63	000000			

Figura 8. Lista de palabras del corpus SAELE versus corpus CEDEL2 (líneas 1-40)

RESULTADOS CUANTITATIVOS DE LOS CORPUS

La aplicación de la herramienta *clusters* en el corpus SAELE se presenta en la figura 9 y revela la existencia de 34 combinaciones formadas por tres palabras, donde *que* aparece en 14 de ellas (líneas 1-6, 13-16, 20-21, 23, 26-27, 29, 31-34).

N	Cluster	Freq.
1	QUE SE LLAMA	30
2	QUE ME GUSTA	25
3	QUE ES UNA	16
4	CREO QUE ES	12
5	QUE ES UN	11
6	ME PARECE QUE	10
7	ME GUSTA MUCHO	9
8	UNA CIUDAD QUE	7
9	UNA PELÍCULA QUE	7
10	Y CREO QUE	7
11	QUE VIENEN DE	6
12	SE TRATA DE	6
13	QUE NO ES	6
14	QUE HAGA HETANEN	6
15	QUE LAS MUJERES	6
16	SIEMPRE SUPI QUE	6
17	ES UNA PELÍCULA	6
18	UN POCO MÁS	6
19	A PESAR DE	6
20	CIUDAD QUE SE	6
21	SUPI QUE HAGA	6
22	ME GUSTA Y	6
23	LOS PAÍSES QUE	6
24	ME GUSTA ES	6
25	Y POR ESO	5
26	LA GENTE QUE	5
27	LAS PERSONAS QUE	5
28	DE LA VIDA	5
29	ES IMPORTANTE QUE	5
30	LOS ESTADOS UNIDOS	5
31	QUE ES MÁS	5
32	QUE SE TRAT/	5
33	NOS DIJO QUE	5
34	QUE ERA UN	5

Figura 9. Lista de agrupaciones de palabras/*clusters*. Corpus SAELE

La figura 10 muestra los resultados de la aplicación de la herramienta *plot/parcelas* al corpus SAELE. Se pueden apreciar el número total de palabras que contiene, las veces en que palabras determinadas aparecen en el texto y en qué proporción por cada mil palabras del corpus.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

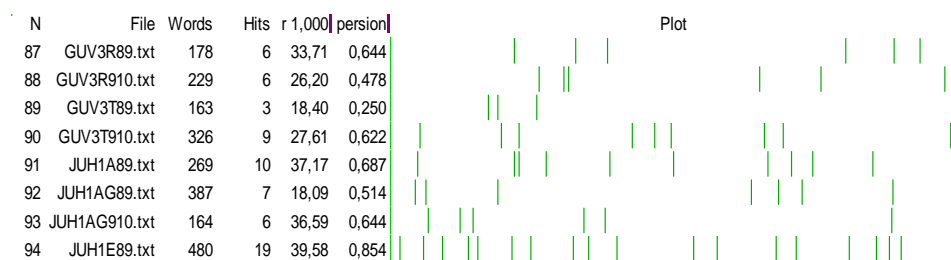


Figura 10. Distribución en contexto de la palabra *que* (líneas 87-94). Corpus SAELE

La herramienta *plot* permitió igualmente detectar que la palabra *que* tuvo una ocurrencia (*hits*) = 27, la más alta de todas en el corpus SAELE (Figura 11).



Figura 11. *Que* en contexto. Corpus SAELE

A continuación en el cuadro 11 presentamos un fragmento del texto que mencionamos con anterioridad donde aparecen marcados 15 *que*⁸⁹.

⁸⁹ En este capítulo marcamos los *que* de los aprendices con un guión bajo (en los ejemplos seleccionados del corpus SAELE).

RESULTADOS CUANTITATIVOS DE LOS CORPUS

Cuadro 11. Ejemplo de un texto argumentativo sobre un artículo periodístico (17 de un total de 27

que)

Según un informe recién realizado por la Organización para la Cooperación y el Desarrollo Económico” (OCDE) México es el país con el nivel más alto de pobreza y desigualdad de ingresos. Entre los miembros de la OCDE México está en el primer puesto en la lista a pesar de *que* el nivel de pobreza y desigualdad ha caído durante los últimos años, pero esta caída no fue suficiente y el país mantiene el primer puesto. Esto también es el caso de la pobreza relativa, aunque ha caído el país está en el primer lugar con mas del 16% de la población viviendo en pobreza, el país está en segundo lugar con unos 22,2% de niños *que* viven en pobreza. La cifra de la pobreza relativa (16%) se puede comparar con países como Suecia, Dinamarca, República Checa y Austria donde menos del 7% de la población vive en pobreza. Según el informe de OCDE Dinamarca y Suecia también son los países con menos desigualdad. Pero en general la desigualdad social ha aumentado en los países desarrollados, particularmente en Alemania, Canadá, Estados Unidos y Noruega. Para suavizar la tendencia de desigualdad en los países desarrollados los gobiernos han intervenido. Redistribuyen los ingresos, fijan más impuestos y aumentan el gasto. El resultado es *que* los gobiernos ahora gastan más en políticas sociales *que* nunca. En la página de OCDE he encontrado dos figuras *que* muestran la situación de la desigualdad y pobreza en México. En la figura 1 se puede ver la evolución de la desigualdad del ingreso medido en el coeficiente de Gini durante los años desde las ochentas. La figura justifica los datos y la información *que* encontrábamos en el artículo. En la figura, se ve muy claro *que* México tiene un coeficiente de Gini mucho más alto *que* el resto de los miembros de OCDE, lo más alto el coeficiente de gini tanto más desigualdad en el país, es decir *que* tienen una desigualdad más alta. También vemos *que* la desigualdad ha caído durante los últimos años, desde los medios de las noventas hasta 2000 la desigualdad ha caído bastante pero a pesar de eso todavía está encima del resto de los miembros de OCDE. Al mismo tiempo vemos *que* la desigualdad de los miembros de OCDE está subiendo, algo *que* puede indicar que dentro de unos años, en largo plazo México puede obtener un nivel de desigualdad *que* el resto de la OCDE. Podemos concluir *que* México está abocado a la dirección correcta mientras *que* el resto de la OCDE está empeorando con una desigualdad subiendo. Es bueno para México pero es muy alarmante *que* la desigualdad en tantos países *que* hay en la OCDE tienen una desigualdad subiendo [...]

La herramienta (*plot*) mostró por otra parte dos ejemplos de textos en que los sujetos habían utilizado *que* solamente una vez (Figura 12).

N	File	Words	Hits	r 1,000	persion	Plot
81	GUV3E910.txt	118	1	8,47	-0,069	
116	JUH2L89.txt	79	1	12,66	-0,069	

Figura 12. *Que* en contexto. Corpus SAELE

En el cuadro 12 presentamos un texto completo donde *que* se utiliza una sola vez.

Cuadro 12. Ejemplo de un texto *Un viaje inolvidable* (1 *que*). Corpus SAELE

Un día en Copenhagen

Tenía once años. Era un día de verano. Mis padres, mi hermano y yo nos levantamos temprano. Cogimos el autobús y fuimos a la ciudad. Allí cogimos otro autobús y fuimos a Copenhagen.

En Copenhagen visitamos unos museos. Por ejemplo un de récords mundiales. Fuimos de compras. Compré una camiseta negra, *que* tengo todavía, pero casi nunca la llevo.

Fuimos a casa más tarde el mismo día. Había sido un día muy interesante

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

En el texto que aparece en el cuadro 13 observamos la aparición de *que* 12 veces.

Cuadro 13. Ejemplo de un texto *Un viaje inolvidable* (12 *que*). Corpus SAELE

Mi primer viaje inolvidable fue en 1975 a Florida.

Los tres hermanos de mi padre vivían allí, después sus retiros. Anteriormente vivían dos de ellos en Nueva York y un vivió en Chicago.

Para mí fue un sueño de mi infancia para ir a los EE.UU y cuando me fui un adulto pude lo realizar.

Recuerdo la mañana en noviembre de 1975, cuando yo y mi novio, más tarde mi marido, dejamos Uddevalla.

Mi padre nos condujo al aeropuerto de Gotemburgo, en el *que* facturamos. Primero tuvimos *que* ir a Copenhague y allí cambiar a un avión más grande, *que* nos traería a Nueva York.

Después un viaje de avión largo estábamos en N.Y, donde tomamos un autobús *que* iba a aeropuertos empalme diferente.

Pronto sentamos en el avión a Florida y en el aeropuerto nos encontraron mi tío y mi tía.

La primera cosa *que* notamos en el aeropuerto fue el calor *que* encontramos, aunque era las nueve por la noche.

El proxima día era el Día de acción de Gracias y fue un gran fiesta de mi tío y mi tía con muchas personas invitados.

Todos los tres tíos vivían cerca unos de otros en el mismo barrio en Clearwater, por eso nos trataron a menudo.

La playa estaba cerca y nos gustaba mucho bañar y nadar, pero un día se advirtió en el altavoz para tiburones en el agua. Era espantoso a oír, pero también nos gustaba estar en espalda y tomar el sol en la arena blanca.

Mi tío, donde estábamos nos llevó a cosas diferentes. Una experiencia divertida fue mirar en carrera de perros, *que* no habíamos visto antes.

El mejor excursion hicimos a Disney World en Orlando, donde fuimos todo el día.

Es un recuerdo *que* nunca olvidaré lo *que* experimenté allí y también y todas personas agradables, *que* nos encontramos.

Después seis semanas dejamos el país muy interesante, para volver a casa a Suecia.

Mis tíos y mis tías son ahora muertos, pero tengo dos primos en los EE.UU, uno en N.Y y uno en Maryland.

Espero *que* alguna vez les visitar o ir a ver el niño del primo, *que* ha comprador una casa en Florida.

Otro ejemplo de frecuencia de uso de *que* aparece en el Cuadro 14. Un texto cuyo tema fue *Mi peli favorita*.

Cuadro 14. Ejemplo de un texto *Mi peli favorita* (11 *que*). Corpus SAELE

Una película *que* me gusta es "Dancer in the dark" de Lars von Trier. En esta película actúan por ejemplo Björk (también una cantante), Peter Stormare (un actor de Suecia) y Katherine Deneuve.

Selma (Björk) es una mujer con un hijo *que* se llama Gene. Los dos alquilan una caravana de Bill y Linda en el campo de Estados Unidos. Bill es una policía y Linda es una ama de casa, los dos son muy amables con Selma y Gene. Bill ha heredado pero no le queda nada de herencia y Linda no lo sabe.

Selma trabaja en una fábrica pero el trabajo es muy difícil para ella porque ve casi nada. Tiene una enfermedad *que* la va a hacer ciega. Sus compañeros no saben que ella tiene esta enfermedad. El hijo tiene la misma enfermedad y Selma aparta todo el dinero *que* gana para una operación de Gene. El hijo no sabe *que* él tiene esta enfermedad. Además Selma tiene un otro trabajo para ganar dinero. Ella hace todo para ayudar a su hijo. A pesar de *que* ella tiene un tiempo duro, siempre está en buen humor y es muy considerado. A menudo sueña *que* ella participa en un musical porque allí no terrible pasa.

Un día cuando Selma va a poner su sueldo entre el dinero ahorrado ella nota *que* el dinero no está. Bill es la única persona *que* sabe donde el dinero está y Selma visita a la casa de Bill. Bill confiesa y pide a Selma *que* no va a decir a la mujer de él, Linda. Pero Selma necesita su dinero y entonces pide Bill a Selma a lo matar. Selma mata Bill y recibe pena capital de muerte.

El fin es muy, muy triste porque se puede ver cuando Selma muere. Es una película muy triste, pero la me gusta mucho. Normalmente no me gusta cuando los personajes cantan en las películas pero en esta, me gusta mucho estas escenas. Björk actúa muy bien y ella recibió Guldpalmen en Cannes por su papel. La película es rodada con una cámara temblada para reflejar la realidad mejor. En sueco se llama esta manera "dogma". Te prometo *que* si ves la película vas a llorar. No se puede mirar esta película sin llorar. Pero la recomiendo.

6.2 Resultados del corpus de control CEDEL2

El corpus de control CEDEL2 (subcorpus de nativos hispanohablantes) está formado de 135 textos producidos por un grupo de nativos hispanohablantes (véase el capítulo 5 para más detalles). El corpus tiene un total de 39.237 palabras y 7.294 tipos de palabras. Los textos fueron escritos por 34 hombres y 91 mujeres.

A continuación aparece la lista de palabras (1-10) más frecuentes del corpus de control CEDEL2 (nativos hispanohablantes) donde aparece *que* en segundo lugar con una frecuencia de 1.550 presente en 134 del total de 135 que contiene el corpus.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

Como se aprecia las cinco palabras más frecuentes son: *de* (F. 2.135), *que* ya mencionada, *la* (F. 1.364), *y* (F. 1.144) y *en* (F. 1.029).

N	Word	Freq.	%	Texts	%
1	DE	2.135	5,44	135	100,00
2	QUE	1.550	3,95	134	99,26
3	LA	1.364	3,48	133	98,52
4	Y	1.144	2,92	135	100,00
5	EN	1.029	2,62	134	99,26
6	EL	1.009	2,57	133	98,52
7	A	896	2,28	135	100,00
8	LOS	517	1,32	119	88,15
9	UN	511	1,30	127	94,07
10	NO	509	1,30	116	85,93

Figura 13. Lista de las 10 palabras más frecuentes del corpus CEDEL2

En el corpus de control CEDEL2 aparece un total de 1.550 concordancias con *que*.

La figura 14 muestra 25 de ellas.

N	Concordance	Set Tag	Word #
1	por el hecho de confrontar el trópico en <i>que</i> vivo con una zona poblada de bajas		42
2	demasiado pero realmente es una zona <i>que</i> debe de ser conocida, con los ojos		162
3	los bosques por parte de los reclusos <i>que</i> dieron origen y vida a la zona hasta		201
4	como castores, truchas y conejos, <i>que</i> se han transformado en		225
5	el equilibrio, y otras acciones <i>que</i> se suman al nefasto calentamiento		240
6	suman al nefasto calentamiento global <i>que</i> , a una hora de vuelo, deterioran día		247
7	deterioran día a día los glaciares. Creo <i>que</i> si hubiera dejado un poco de tiempo		260
8	al regreso, al enorme temor por lo <i>que</i> ya parece ser un proceso		290
9	mode el progresismo de tres al cuarto <i>que</i> se preocupa por el entorno, los		46
10	y un mundo mejor, pero el hecho es <i>que</i> se siguen marginando temas mucho		64
11	temas mucho más importantes <i>que</i> el que yo me fume cuatro Camel al		72
12	temas mucho más importantes que el <i>que</i> yo me fume cuatro Camel al día.		74
13	¿con qué derecho se cree el Estado <i>que</i> puede llamarnos imbéciles a la cara		109
14	tabaco? Es bastante triste, y más aún <i>que</i> un fumador tenga que ser un		132
15	triste, y más aún que un fumador tenga <i>que</i> ser un apestado. Entre ser		136
16	mi bolsillo. Es simplemente deleznable <i>que</i> nos hagan aprender, respirar y fumar		182
17	nos hagan aprender, respirar y fumar lo <i>que</i> luego quieren quitarnos a tortazos.		190
18	Una película <i>que</i> me ha llamado la atención		2
19	República Democrática Alemana al <i>que</i> se le encarga la misión de vigilar a		37
20	pero una (entre varias) razones por la <i>que</i> deciden vigilarle es porque el		67
21	a punto de pillarlo, porque la actriz, <i>que</i> después de un desplante al ministro		141
22	el agente de la policía secreta <i>que</i> estaba encargado de la vigilancia		180
23	deshacerse de las pruebas antes de <i>que</i> los agentes de la policía secreta que		200
24	que los agentes de la policía secreta <i>que</i> registran la casa del escritor las		207
25	encuentren. En ese momento, la actriz, <i>que</i> está presente en el registro,		220

Figura 14. Concordancias con *que*. Corpus de control CEDEL2

RESULTADOS CUANTITATIVOS DE LOS CORPUS

El resultado de la búsqueda de las colocaciones más frecuentes con *que* en el corpus de control CEDEL2 aparece en la figura 15. Se aprecia a *de que* (línea 2) y *lo que* (línea 13).

N	Word	With elation	Texts	Total
1	QUE	que 0,000	134	1.904
2	DE	que 0,000	125	649
3	LA	que 0,000	118	479
4	Y	que 0,000	113	366
5	EL	que 0,000	112	355
6	EN	que 0,000	111	348
7	A	que 0,000	110	337
8	NO	que 0,000	91	258
9	ES	que 0,000	94	255
10	SE	que 0,000	95	239
11	LOS	que 0,000	86	204
12	UNA	que 0,000	85	183
13	LO	que 0,000	73	176
14	UN	que 0,000	85	173
15	POR	que 0,000	73	146
16	MÁS	que 0,000	58	121
17	CON	que 0,000	67	118
18	ME	que 0,000	52	116
19	LAS	que 0,000	56	109
20	DEL	que 0,000	62	100
21	PARA	que 0,000	60	92
22	AL	que 0,000	56	92
23	SU	que 0,000	38	74
24	HAY	que 0,000	33	68
25	PERO	que 0,000	43	66

Figura 15. Colocaciones con *que*. Corpus de control CEDEL2

El cuadro 15 muestra el número total de concordancias con las combinaciones de palabras con *que* objetos de la presente investigación. Se aprecia que la combinación *lo que* encabeza con 82 concordancias, *algo que* aparece exclusivamente en 6 concordancias y *dice que* no aparece utilizada en absoluto lo que contrasta con el número de apariciones en el corpus de aprendices sueco SAELE.

Cuadro 15. Concordancias de combinaciones de palabras con *que*. Corpus de control CEDEL2

Concordancias	Lo que	De que	Algo que	Dice que
Total	82	62	6	0

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

En la figura 16 aparecen 25 de las 82 concordancias con *lo que* en el Corpus de control CEDEL2.

N Concordance	Set Tag	Word #	t. #	os. a. #	os. d. #	os. t. #	os.	File	%
1		paso, al regreso, al enorme temor por lo que ya parece ser un proceso	289	8 8%	0 7%	0 7%		A2.txt	97%
2		nos hagan aprender, respirar y fumar lo que luego quieren quitarnos a	189	6 7%	0 6%	0 6%		AC.txt	96%
3		otra ciudad de EEUU. Creo que lo que le hace tan interesante es la	24	3 4%	0 7%	0 7%		APP.txt	7%
4		a casa de mis abuelos, ¡No reconocía lo que veía! ¿Qué sucedía? ¿El mar se	130	6 6%	0 0%	0 0%		APR.txt	50%
5		puede definir tan escurridizo concepto. Lo que se tiene a los homosexuales es,	290	8 4%	0 1%	0 1%		AR.txt	39%
6		todas las razas, culturas y religiones y lo que más me sorprende es que todo el	135	7 9%	0 2%	0 2%		BRV.txt	22%
7		India, son dos barrios de la ciudad. Lo que más me ha gustado ha sido el	206	11 4%	0 3%	0 3%		BRV.txt	34%
8		mi principal objetivo cuando vine aquí. Lo que menos me gusta de Toronto es el	291	15 6%	0 7%	0 7%		BRV.txt	48%
9		y todavía están quitando escombros y lo que más me llamó la atención fue una	506	23 4%	0 2%	0 2%		BRV.txt	82%
10		bastante bien el viaje desde España, lo que agilizó ciertas cosas de la	62	3 3%	0 6%	0 6%		C.txt	61%
11		intentan ir más allá, ver más allá de lo que pueden ver nuestros sentidos.	355	9 2%	0 0%	0 0%		CMH.txt	78%
12		y por ello creo que me será más fácil. Lo que más destaco de la película es,	43	2 5%	0 7%	0 7%		CMM.txt	7%
13		es en la infancia cuando se "descubre" lo que es una familia, pero se olvida a	167	5 1%	0 7%	0 7%		CMM.txt	27%
14		sufren un gran choque emocional por lo que les ocurre son las mujeres: la	367	9 5%	0 9%	0 9%		CMM.txt	59%
15		que viajar por todo el mundo y conocer lo que otras culturas pueden ofrecerse.	108	3 7%	0 4%	0 4%		DVC.txt	74%
16		al tema de adoptar niños, creo que lo que un niño necesita es crecer en un	171	9 2%	0 5%	0 5%		EGB.txt	67%
17		eso no es un escándalo. En mi opinión lo que un niño necesita es que le	236	14 6%	0 9%	0 9%		EGB.txt	90%
18		del otro mundo, supongo que es con lo que sueña cualquier persona con un	143	7 7%	0 9%	0 9%		EH.txt	72%
19		madura, muy consciente de lo que desean y las razones por las que	63	3 2%	0 1%	0 1%		EHTT.txt	22%
20		las cosas son bastante diferentes a lo que él esperaba y la verdadera razón	149	6 5%	0 9%	0 9%		EMP.txt	68%
21		a medida que avanza el relato, de lo que podría parecer al principio, pues	187	4 7%	0 3%	0 3%		ENB.txt	62%
22		gracias a los glaciares que aplanaron lo que pudo haber habido, hermoso lago	29	1 2%	0 9%	0 9%		ESM62.txt	10%
23		millonarios y celebridades. Era, lo que en la jerga técnica se	418	16 8%	0 8%	0 8%		FJ.txt	68%
24		embargo, el Sur es bastante plano, por lo que hay una diferencia de varios	27	1 0%	0 6%	0 6%		HNN.txt	5%
25		anotamos un punto a su favor en lo que respecta al transporte público:	105	5 6%	0 1%	0 1%		HRSB.txt	31%

Figura 16. Concordancias con *lo que*. Corpus de control CEDEL2

La combinación o agrupación de tres palabras *por lo que* (Figura 17) es empleada con frecuencia por los nativos hispanohablantes, pero también otras preposiciones como *de* o *en*. En el uso de esta combinación de palabras se destacan expresiones idiomáticas como: dar por hecho lo que no está hecho; ver más allá de lo que puedan ver y hacer lo que quiera. También observamos a *lo que* en su papel de conector textual.

RESULTADOS CUANTITATIVOS DE LOS CORPUS

N Concordance

2 teniendo todos los días fiebre **por lo que** no me podían dar el alta y sin que

3 a la docencia universitaria, **por lo que** me decidí a dar el primer paso,

4 sufren un gran choque emocional **por lo que** les ocurre son las mujeres: la

5 niño es una tarea bastante difícil, **por lo que** en la mayoría de los casos, las

6 llegar a ser profesora universitaria, **por lo que** no dejaré mi labor investigadora,

7 absurdo que consiste en dar **por hecho lo que** no está hecho, en dar por

8 poco mejor, no te hacen contrato, **por lo que** no puedes acreditar luego tu

9 paso, al regreso, al enorme temor **por lo que** ya parece ser un proceso

10 que no está hecho, en dar **por sucedido lo que** no ha sucedido. Es evidente que,

11 como debiera, y muy frío, **por lo que** abundan abrigos de piel, y

12 madura, muy consciente de **lo que** desean y las razones por las que

13 Stewart, que vive en Nueva Zelanda, **por lo que** ella y su hija Flora tienen que

14 embargo, el Sur es bastante plano, **por lo que** hay una diferencia de varios

15 a su tendencia sexual) que lo aparte de **lo que**, **por** naturaleza, se supone

Figura 17. Combinación de palabras *por lo que*. Corpus SAELE

De que es la combinación de palabras que ocupa el segundo lugar en el corpus CEDEL2 en cuanto al número total de concordancias. La figura 18 muestra 25 de esas concordancias. Las expresiones idiomáticas *darse cuenta de que* (líneas 2, 10, 13, 15 y 25) y *el hecho de que* (líneas 7 y 18) son de las más utilizadas por los nativos hispanohablantes.

N Concordance	Set Tag	Word #	t. #	os. a. #	os. d. #	os. t. #	os.	File	%
1 logra deshacerse de las pruebas antes de que los agentes de la policía secreta		199	6 5%	0 0%		0 0%		ACP.txt	61%
2 Después de leerlos, se da cuenta de que había sido el agente el que le		272	9 5%	0 2%		0 2%		ACP.txt	83%
3 hace ya un par de años con el doble fin de que las personas no fumadoras no		18	0 1%	0 5%		0 5%		AJSL-.txt	4%
4 así como incrementar las posibilidades de que los fumadores dejaran de fumar.		38	0 6%	0 1%		0 1%		AJSL-.txt	11%
5 A mí me parece muy buena la idea de que no se pueda fumar en lugares		8	0 0%	0 6%		0 6%		AOC.txt	4%
6 a muchos. A ellos se les culpa de que nadie pueda comprarse una casa		295	21 7%	0 9%		0 9%		APP.txt	90%
7 en cuando nace un ortodoxo. El hecho de que exista y más se use una voz		144	5 1%	0 0%		0 0%		AR.txt	19%
8 trayectoria de pareja estable en fin de que el hijo pueda desarrollarse con la		696	15 7%	0 7%		0 7%		AR.txt	97%
9 Cuando los padres de ella se enteran de que anda con este chico la prohíben		73	5 5%	0 2%		0 2%		CAC.txt	31%
10 ambos se reconocen, se dan cuenta de que siguen enamorados y comienzan		142	7 0%	0 1%		0 1%		CAC.txt	60%
11 a escondidas. El príncipe sospecha de que su prometida le es infiel y manda		155	8 5%	0 7%		0 7%		CAC.txt	66%
12 tribu. La versión más aceptada es la de que esta civilización desapareció		116	4 1%	0 7%		0 7%		CLB.txt	35%
13 que lo traicionó, el padre se da cuenta de que su socio lo ha engañado y lo ha		300	8 1%	0 9%		0 9%		CMM.txt	49%
14 con la única diferencia, claro está, de que algunos no optarán por la filología		195	9 1%	0 5%		0 5%		EH.txt	95%
15 experiencia me ha hecho darme cuenta de que el país merece una visita, o dos,		103	6 0%	0 2%		0 2%		EHT.txt	31%
16 sin perjuicio de todas los instrumentos de que dispone el estado de derecho		329	9 1%	0 2%		0 2%		ELL.txt	92%
17 Des de que nació vivo en Barcelona, es una		1	0 3%	0 0%		0 0%		ICL.txt	0%
18 no lo quieran ver por el simple hecho de que seamos catalanes y de que		197	10 9%	0 5%		0 5%		ICL.txt	64%
19 hecho de que seamos catalanes y de que defendamos la pluralidad. La		202	10 2%	0 7%		0 7%		ICL.txt	66%
20 al rescate de la chica, justo antes de que el coche la atropellara. Ellos		132	6 9%	0 0%		0 0%		JDG.txt	57%
21 con la bella doncella, a pesar de que ambos estaban comprometidos		189	10 7%	0 6%		0 6%		JDG.txt	85%
22 de los casos mezquino. El triste tópico de que en el extranjero se te valora más		114	6 7%	0 9%		0 9%		JEPP.txt	40%
23 Después de dos días allí me di cuenta de que el dolor de espalda persistía		155	7 7%	0 7%		0 7%		JRG.txt	37%
24 pruebas llegaron a la conclusión de que , por circunstancias que aún no		334	14 6%	0 9%		0 9%		JRG.txt	80%
25 Estoy de acuerdo con la idea básica de que nuestra constitución y todo el		59	2 6%	0 0%		0 0%		LMA.txt	20%

Figura 18. Concordancias *de que* (25 de un total de 62), Corpus de control CEDEL2.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

La combinación menos utilizada por los nativos hispanohablantes es *algo que* (6). Obsérvese en la figura 19 que a la misma le precede el verbo ser en 5 de las 6 líneas de concordancia.

NConcordance		Set	Tag	Word #	t. #	os. a. #	os. d. #	os. t. #	os.	File	%
1	El problema del terrorismo es <i>algo que</i> a los españoles nos afecta	6	0	1%	0	1%	0	1%	0	VLS.txt	1%
2	tampoco me parece una cosa que sea <i>algo que</i> se debe hacer, porque hay que	99	4	8%	0	2%	0	2%	0	NJF.txt	43%
3	españolas hace varias décadas, <i>algo que</i> puede afirmarse es que se	165	7	2%	0	1%	0	1%	0	MVGA.txt	60%
4	de sangre y tiendas destrozadas, es <i>algo que</i> siempre tendré en mi recuerdo.	384	15	5%	0	4%	0	4%	0	MERM.txt	85%
5	Toda esa variedad... pero que hay <i>algo que</i> une a todos, y que hace que se	172	7	6%	0	6%	0	6%	0	BMN.txt	54%
6	cualquiera, porque la homosexualidad es <i>algo que</i> los homosexuales sienten	569	13	2%	0	0%	0	0%	0	AR.txt	79%

Figura 19. Concordancias con *algo que* (6 de un total de 6). Corpus de control CEDEL2

No aparecen concordancias con la combinación *dice que* en el corpus de control CEDEL2.

La figura 20 muestra el resultado de la aplicación de la herramienta denominada *Keywords* lista, a la que se contraponen la lista de palabras formadas del corpus CEDEL2 al SAELE para ver qué palabras son utilizadas en ambos corpus y en cuál frecuencia aparecen.

RESULTADOS CUANTITATIVOS DE LOS CORPUS

N	Key word	Freq.	%	. Freq.	RC.	%	eyness]	P	emmas]	Set
1	DE	2.135	5,44	1.508	4,07	79,63	000000]			
2	QUE	1.550	3,95	1.094	2,95	57,30	000000]			
3	A	896	2,28	613	1,65	39,34	000000]			
4	NI	44	0,11	4	0,01	36,80	000000]			
5	PAREJAS	27	0,07	0		35,93	000001]			
6	VIVO	36	0,09	2		35,13	000006]			
7	HOMOSEXUALES	25	0,06	0		33,27	000051]			
8	REGIÓN	22	0,06	0		29,28	000598]			
9	LO	257	0,65	141	0,38	28,20	001066]			
10	ZONA	21	0,05	0		27,94	001219]			
11	O	176	0,45	85	0,23	27,55	001501]			
12	AL	252	0,64	140	0,38	26,55	002533]			
13	TE	31	0,08	3		25,29	004901]			
14	PLANES	19	0,05	0		25,28	004921]			
15	PASO	19	0,05	0		25,28	004921]			
16	SE	464	1,18	307	0,83	24,16	008830]			
17	LOS	517	1,32	650	1,75	-24,03	009463]			
18	MAS	6	0,02	35	0,09	-24,39	007845]			
19	TIEMPO	53	0,14	111	0,30	-24,43	007670]			
20	LIBRE	7	0,02	38	0,10	-25,29	004906]			
21	PERO	182	0,46	278	0,75	-26,15	003134]			
22	PAÍSES	12	0,03	49	0,13	-26,24	002989]			
23	TIENE	42	0,11	100	0,27	-27,85	001283]			
24	FUIMOS	11	0,03	49	0,13	-28,23	001048]			
25	MUJERES	12	0,03	51	0,14	-28,27	001027]			
26	LLAMA	7	0,02	41	0,11	-28,64	000841]			
27	DICE	6	0,02	40	0,11	-30,13	000376]			
28	EJEMPLO	7	0,02	44	0,12	-32,06	000120]			
29	ES	461	1,17	619	1,67	-33,49	000042]			
30	ESPAÑA	34	0,09	99	0,27	-37,02	000000]			
31	MUCHO	56	0,14	152	0,41	-51,77	000000]			
32	ERA	67	0,17	173	0,47	-54,84	000000]			
33	CUANDO	85	0,22	205	0,55	-58,43	000000]			
34	MUY	128	0,33	295	0,80	-77,98	000000]			
35	TAMBIÉN	54	0,14	182	0,49	-80,98	000000]			
36	GUSTA	14	0,04	120	0,32	-102,30	000000]			
37	#	110	0,28	325	0,88	-124,29	000000]			

Figura 20. Contraposición de las listas de palabras formadas según el *Keywords* lista. CEDEL2 versus SAELE

La aplicación de la herramienta *clusters* al corpus de control CEDEL2 reveló la existencia de un total de 47 agrupaciones de palabras. La palabra *que* aparece en mayor índice de frecuencia en la agrupación *verdad es que* (Figura 21, línea 2).

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

N	Cluster	Freq.
1	LA VERDAD ES	18
2	VERDAD ES QUE	16
3	EN EL QUE	16
4	CUENTA DE QUE	13
5	QUE HE VISTO	12
6	POR LO QUE	12
7	PELÍCULA QUE HE	10
8	EN LA QUE	10
9	DE LO QUE	10
10	Y CREO QUE	9
11	QUE NO ES	9
12	ÚLTIMA PELÍCULA QUE	9
13	LA ÚLTIMA PELÍCULA	9
14	QUE EN EL	9
15	QUE NO ME	9
16	NO HAY QUE	9
17	QUE NO SE	9
18	QUE MÁS ME	8
19	DE QUE EL	8
20	QUE EN LA	7
21	VA A SER	7
22	LO ÚNICO QUE	7
23	CON LO QUE	7
24	TODO EL MUNDO	7
25	DE LOS QUE	7
26	LA GENTE QUE	7
27	QUE ME HA	6
28	LO QUE NO	6
29	QUE LO QUE	6
30	ES VERDAD QUE	6
31	QUE LAS PAREJAS	6
32	QUE VER CON	6
33	QUE VA A	6
34	QUE OLVIDAR QUE	6
35	HAY QUE OLVIDAR	6
36	HECHO DE QUE	6
37	QUE HAY GENTE	6
38	LO QUE SE	6
39	A LOS QUE	6
40	EL HECHO DE	6
41	A PESAR DE	6
42	UNA VEZ QUE	6
43	ES QUE LA	6
44	HAY GENTE QUE	6
45	Y LO QUE	5
46	VIDA Y QUE	5
47	UNA PELÍCULA QUE	5
48	A MEDIDA QUE	5
49	PESAR DE QUE	5
50	DE QUE NO	5
51	EL QUE NO	5
52	LO QUE MÁS	5
53	MÁS QUE UNA	5
54	LO QUE LE	5
55	QUE TODO EL	5
56	QUE UN NIÑO	5
57	DE AQUELLOS QUE	5
58	QUE ES LA	5
59	DE LA PELÍCULA	5

RESULTADOS CUANTITATIVOS DE LOS CORPUS

Figura 21. Lista de agrupaciones de palabras/*clusters*. Corpus de control CEDEL2

La aplicación de la herramienta *plot* refleja visualmente, como se aprecia en la figura 22, las veces que se utiliza *que* en quince textos compilados del corpus de control CEDEL2.

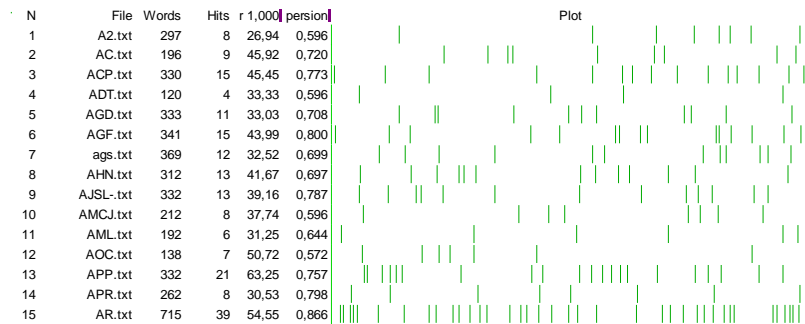


Figura 22. Distribución de *que* en los textos. Corpus de control CEDEL2

6.3 Resultados del corpus de control ARGUS

El corpus de control ARGUS (nativos suecos) elegido está formado de 135 textos escritos por 73 mujeres y 52 varones. Tiene un total de palabras de 77.650 (Figura 23).

N	Overall
text file	Overall
file size	452.150
tokens (running words) in text	77.650
tokens used for word list	76.694
sum of entries	
types (distinct words)	7.819

Figura 23. Información general. Corpus de control ARGUS

En la Figura 24 se aprecia que la palabra más frecuentemente empleada en este corpus fue *att/que* (con una frecuencia 2.908 = 3,75%), apareciendo además en el total de los 135 textos que componen este corpus. Le siguen *och/y* (2.466 = 3,18%), *det/el/la/lo* (2.049 = 2,64%), *i/en/dentro de* (1.745 = 2,25%), *är/ser o estar* (1.713 = 2,21%) y *som/que/como* (1.652 = 2,13%).

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

N	Word	Freq.	%	Texts	%
1	ATT	2.908	3,75	135	100,00
2	OCH	2.466	3,18	135	100,00
3	DET	2.049	2,64	134	99,26
4	I	1.745	2,25	135	100,00
5	ÄR	1.713	2,21	135	100,00
6	SOM	1.652	2,13	135	100,00
7	FÖR	1.277	1,64	135	100,00
8	JAG	1.264	1,63	132	97,78
9	EN	1.162	1,50	135	100,00
10	DE	1.003	1,29	128	94,81
11	PÅ	990	1,27	135	100,00
12	#	956	1,23	121	89,63
13	INTE	910	1,17	134	99,26
14	BARN	893	1,15	112	82,96
15	MAN	831	1,07	123	91,11
16	OM	799	1,03	132	97,78
17	MED	770	0,99	132	97,78
18	TILL	732	0,94	135	100,00
19	AV	704	0,91	128	94,81
20	HAR	697	0,90	133	98,52
21	SÅ	641	0,83	128	94,81
22	KAN	625	0,80	126	93,33
23	VI	575	0,74	105	77,78
24	ETT	571	0,74	128	94,81
25	SIG	463	0,60	125	92,59

Figura 24. Lista de palabras más frecuentes (1-25). Corpus de control ARGUS

El equivalente de *que*, es decir la palabra sueca *att* aparece en 134 concordancias, de las cuales se muestran 30 ejemplos en la figura 25.

RESULTADOS CUANTITATIVOS DE LOS CORPUS

N Concordance

1 från böckerna. Lilla Puh lärde oss att vara omtänk[samma och inte bara
2 glada starka flickan visade oss barn att inte se så allvarligt på allt, utan vi ska
3 Jag kan då i alla fall aldrig minnas att jag sett Pippi ledsen, oavsett vilken
4 Av lilla Emils busiga liv fick vi reda på att lite bus var okej, men ibland måste vi
5 kan det då vara möjligt för ett litet barn att förstå sådana resonemang vid fem till
6 av det. De berättar ofta för barn för att vidga deras erfarenhets- och
7 förståelse för sammanhang i den". Att författare arbetar medvetet med
8 Rimsten-Nilsson har rätt i. Det är viktigt att anpassa språket till hur mycket den
9 Däremot har de fått lära sig från början att om man är stark som "Starke Adolf",
10 Språket i en bambok är också viktigt för att utvidga barnets ordförråd. Som vi kan
11 bok om "Bamböcker och läslust" är att lärare upptäcker att elever numera allt
12 och läslust" är att lärare upptäcker att elever numera allt oftare saknar
13 när det kommer till skolan, och att de under skoltiden inte utvecklar ett
14 ett tillräckligt avancerat språk för att kunna tillgodogöra sig
15 upptar all barnens fritid, det är bara att titta på sina yngre syskon. Det är inte
16 jag bara komma på en idé, och det är att barnens lärare ger dem fler böcker i
17 skrev i Bambarometern (98/99) att andelen bokläsare bland de yngsta
18 (Barn böcker och läslust, 1981) skriver att seriösa bamboksförfattare arbetar
19 speciellt med språket i sina böcker för att främja barnens utveckling.
20 i kring barnens egna liv och hjälper dem att själva förklara vad som
21 mer förenklat sätt. Förstår inte föräldrar att ett barn behöver läsa? Kanske är det
22 i de tidiga åren i ett barns liv som gör att många barn saknar empati och andra
23 lägre åldrama? Jag vet : det fanns ingen att läsa för dessa ungdomar när de var
24 de var små. Rimsten-Nilsson förklarar att författarna använder ett språk, i sina
25 erfarenhet, vilket inte bara hjälper barn att få ord till att beskriva egna
26 vilket inte bara hjälper barn att få ord till att beskriva egna upplevelser utan också
27 upplevelser utan också hjälper barn att se ett samband i vad de upplever.
28 de upplever. Orsaker och följder. Tror ni att en 9-årig snattare förstår följderna och
29 och fantasieggande och ledde till att jag själv försökte skriva egna böcker.
30 barn detta minnet som kan komma att grunda barnets framtid. Kerstin

Figura 25. Concordancias con *att*. Corpus de control ARGUS

En la figura 26 se muestra el resultado de la aplicación de la herramienta *clusters*. La palabra *att* aparece en diferentes posiciones en 31 de las 35 agrupaciones de palabras.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

N	Cluster	Freq.	Set	Length
1	ATT DET äR	116		3
2	JAG TROR ATT	38		3
3	JAG TYCKER ATT	35		3
4	ATT DET FINNS	30		3
5	FÖR ATT FÅ	29		3
6	ATT DET VAR	28		3
7	TYCKER ATT DET	27		3
8	JAG ANSER ATT	26		3
9	FÖR ATT KUNNA	24		3
10	ATT DET INTE	24		3
11	DET äR VIKTIGT	23		3
12	ISTÄLLET FÖR ATT	20		3
13	ATT PRATA OM	19		3
14	TYCKER JAG ATT	19		3
15	äR VIKTIGT ATT	18		3
16	SE TILL ATT	18		3
17	TILL ATT LÄSA	17		3
18	OCH DET äR	16		3
19	JAG HOPPAS ATT	16		3
20	KAN GE BARN	16		3
21	ATT LÄSA FÖR	15		3
22	ATT LÄSA BÖCKER	14		3
23	ATT LÄRA SIG	14		3
24	ATT MAN SKA	14		3
25	ATT LÄSA BARNBÖCKER	14		3
26	ANLEDNING ATT PRATA	14		3
27	ATT DE INTE	13		3
28	äR DET VIKTIGT	12		3
29	JAG ATT DET	12		3
30	TYCKTE ATT DET	12		3
31	ATT VI SKA	12		3
32	ATT MAN INTE	12		3
33	TROR ATT DET	12		3
34	ATT SE TILL	11		3
35	BARNEN ATT LÄSA	11		3

Figura 26. Agrupaciones de palabras/*clusters* con *att* (35 de un total de 305). Corpus de control

ARGUS

El pronombre sueco *som*, que en algunos casos puede ser equivalente a *que* en lengua española aparece en 1652 concordancias del corpus de control ARGUS. La figura 27 muestra algunas de las combinaciones de palabras con *som*, equivalentes a las combinaciones con *que* que son objeto de la presente investigación, como en los ejemplos: *lo que/det som* (línea 2); *lo que/vad som* (línea 11) y *algo que/något som* (línea 8).

RESULTADOS CUANTITATIVOS DE LOS CORPUS

N Concordance

1 med språket. De använder ett språk som är förankrat i barnens språk och
2 /.../ De berättar om orsaker till det som händer och följderna av det. De
3 bamböcker, då främst Astrid Lindgrens, som en symbol för en riktigt stark man.
4 har ofta en "anabolauppumpad man" som symbol för någon riktigt stark. Men
5 lära sig från början att om man är stark som "Starke Adolf", då är man riktigt
6 viktigt för att utvidga barnets ordförråd. Som vi kan läsa i Rimsten-Nilssons bok
7 utvecklas förrådet med abstrakta ting som känslor, förhållande, regler för rätt
8 regler för rätt och orätt mm. Något som vi tråkigt nog också kan läsa i
9 böcker i hemläxa. Har inte ni någon idé som ni kan dela med er av i en
10 få igång läsandet bland barn igen. För som sagt, "bättre sent än aldrig".
11 och hjälper dem att själva förklara vad som känslomässigt pågår i deras liv, på
12 av böcker i de tidiga åren i ett barns liv som gör att många barn saknar empati
13 och andra normer och värderingar som vi "vuxna" kan anse vara normala.
14 använder ett språk, i sina böcker, som är förankrat i barnens språk och
15 alla ihåg Lotta på Bråkmakargatan, som kan allt, Nicke Nyfiken, Madicken,
16 böcker från Disney-klubben? Jag säger som många före mig: det var bättre förr.
17 eller pappa läsa den där favoritboken som man efter ett tag kunde utantill. Vill
18 föräldrar unna sina barn detta minnet som kan komma att grunda barnets
19 och upp till gymnasiet vittnar om elever som nu saknar ett "elementärt verbalt
20 arbetsgivare. Slutsatsen kan vem som helst dra: Barn behöver läsa. På
21 mycket om den allt mindre tid som barn ägnar åt läsning, vad som kan
22 tid som barn ägnar åt läsning, vad som kan göras åt det och hur det
23 språkutveckling. Jag har hört åsikter som att barns språkutveckling inte störs
24 hävdar att det är viktigare att man som vuxen och förälder pratar mycket
25 pratar mycket med barnen om vad som har hänt under dagen osv. Att det
26 bokens betydelse kommer in i bilden. Som ett verktyg till utvecklande samtal.
27 medvetet med språk och handling som ska lära, hjälpa och utveckla
28 skriver om händelser som barn känner igen sig i, och i och
29 att det finns en person i en berättelse som barnet identifierar sig med skriver
30 ger ett tillfälle att prata om något som annars hade varit svårt för barnet.

Figura 27. Concordancias con *som*. Corpus de control ARGUS

Las figuras 28 y 29 muestran respectivamente 33 concordancias (de un total de 33) para la combinación *något som/ algo que* en la figura 28 y 14 concordancias (de un total de 14) con *någon som/ algo que* en la figura 29. Ambas son combinaciones o palabras en sueco equivalentes a las de la lengua española⁹⁰ cuyo uso y función son objeto de análisis en el presente estudio.

⁹⁰ Para más detalles sobre los equivalentes suecos de las combinaciones de palabras con *que* véase el capítulo 4.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

N Concordance

1 Ungdomar duar varandra, "ni" är något som används för att visa respekt
2 något en gång för alla? Något drastiskt. Något som märks. Den främsta åtgärd
3 dess invånare måste tillsammans göra något som märks för att minska
4 En svordom är inte något fullt utan något som sätter extra kraft bakom
5 dåliga. Men väger "ja" tyngre än "nej"? Något som oroar de flesta är om
6 finns runt omkring oss. Det är något som framför åsikter eller fakta till
7 med sin familj eller skaffa sig en hobby! Något som jag tycker är skrämmande är
8 former. Och anorexia är inte längre något som drabbar få utan en vanligt
9 här hos oss. Barnarbetet är något som till slut måste förkastas och
10 leder tyvärr till svält och senare död. Något som inte är helt ovanligt är
11 Jag tänkte avsluta mitt tal här idag med något som jag tycker är väldigt konstigt,
12 det som utmärker en svordom jo, det är något som är oacceptabelt och om de
13 och kan betraktas som själva giftet, något som man skall helst akta sig för.
14 är ganska enkelt antingen har personen något som dom andra inte har exempel
15 att man känner sig kritisk. Sexualitet är något som alla på något sätt upplever,
16 lyssnar på barnarbetarna, ingen gör något som förbättrar deras vilkor eller
17 som tex vår fackliga trygghet. Det är något som inte existerar i barnslavarnas
18 är fullt i språket. Va och andra små är något som retar människor. De småord
19 för personens framtida tänkande, är något som föräldrar borde vara mer
20 de i en fram när barnet läser själv. Något som inspirerar till att läsa böcker
21 och bli lika duktig som denna någon. Något som utvecklas mycket när man
22 skolpojkar, eller ilska tjuvar. Något som jag beundrar hos henne är att
23 det något roligt och inspirerande, inte något som man tvingas till. För jag säger
24 det vara lättare för dem att prata om något som de till exempel är ledsna för
25 och barnboks författare ska det inte vara något som helst problem att finna
26 man gjorde något betydelsefullt. Det är något som jag tycker alla barn borde få
27 Jag anser därför att detta är något som man inte har rätten att neka
28 struntade i en. Böcker var inte alls något som man intresserade sig för, utan
29 Svenska. Ta Rinkeby, där pratar man något som de kallar Rinkebysvenska.
30 Detta tror jag är viktigt, för fantasin är något som alla använder mycket ofta.
31 Böckerna ger ett tillfälle att prata om något som annars hade varit svårt för
32 hade varit svårt för barnet. Detta är något som jag själv hade stor glädje av
33 förhållande, regler för rätt och orätt mm. Något som vi tråkigt nog också kan läsa

Figura 28. Concordancias con *något som*. Corpus de control ARGUS

N Concordance

1 tycker jag är ett sätt av visa respekt för någon som är äldre. När jag niar, så niar
2 vara självklart, man ska inte tvinga någon som är nästan vuxen att lära sig
3 man inte få en förmanande blick och någon som säger: Det heter faktiskt
4 och brottsligheten ökar, när det inte är någon som har tid för de som ska skapa
5 leda till dödsfall. Vad är det för liv för någon som är så ung? Vi kan ändra på
6 och tidsödande syssloma i hemmet. Någon som ser till att det finns mat på
7 när man kommer hem från jobbet. Någon som ser till att man slipper städa
8 på att anställa en hushållerska. Någon som kan laga mat och städa.
9 Någon som kan laga mat och städa. Någon som ser till att man inte behöver
10 inte nej till. Jag kan bara inte förstå hur någon som använder barn till arbete kan
11 vuxna" Det är precis som bamen läser någon som förstår sig på ens egna
12 barn läser böcker så får de en förebild, någon som de kan efterlikna. "Ibland är
13 för barnet, också i tidig ålder med någon som läser för dem. De utvecklas
14 ett enkelt samtalsämne är ju böcker, någon som de kanske har läst eller fått

Figura 29. Concordancias con *någon som*. Corpus de control ARGUS

RESULTADOS CUANTITATIVOS DE LOS CORPUS

La figura 30 muestra 30 de las 48 concordancias con *det som*, combinación que puede equivaler a *que* y a *lo que* en determinados contextos fraseológicos.

N Concordance

1 /.../ De berättar om orsaker till **det som** händer och följderna av det. De
2 finns inom synhåll för barnet och ord för **det som** är abstrakt tex. känslor kommer
3 det verkligen hela sanningen? Nej, för **det som** gjort oss till de vi är idag, är vår
4 anpassa oss. Och vad kan inte ge allt **det som** ett barn behöver om inte
5 inte gärna frågade om. Den biten, av **det som** läsandet ger, är otroligt viktig för
6 läst godnattssagan innan sovdags? **Det som** sker idag är farligt!
7 ett liknande problem när han var ung. **Det som** skilde sig var att han inte var
8 eget skapande. Jag är övertygad om att **det som** Kerstin Rimsten-Nilsson säger
9 vilja läsa. Men vad är roligt för dem? Jo **det som** existerar i deras värld. Kerstin
10 fantasivärd, genom att de kan benämna **det som** finns i deras yttre värld. Boken
11 blir det att man talar om det. Och det är **det som** Kerstin Rimsten-Nilsson skriver
12 dagis, fritids, lekis eller i skolan. Och **det som** är lite roligt är att barn/böcker
13 ofta spontana kommentarer och då är **det som** en vuxen lätt att utveckla
14 mer utvecklat språk får vi. Det är just **det som** är så bra med litteraturen, att
15 det är att läsa bamböcker. Tyvärr verkar **det som** intresset för att läsa bamböcker
16 sväva sig bort i fantasin och leva sig in i **det som** hände i boken. Bara försvinna
17 förändras efter livets alla olika villkor. **Det som** är populära uttryck i ungdomen
18 på vissa saker han gjorde och det var ju **det som** gjorde böckerna så bra. Nu när
19 ta ut svängarna", skriver Petter. Så **det som** han en gång hatade har idag
20 FÖR BARNBOKEN! Vad är **det som** har orsakat att dagens föräldrar
21 på vårt tankesätt i det fortsatta livet. **Det som** jag nämt ovan är bara det allra
22 Jag skulle vilja ställa frågan vad är **det som** gör att läsandet går ner? Jag
23 en krydda i språket eller anses **det som** stötande? Åsikterna är olika om
24 att de språkliga argumenten är att; "**Det som** är för tycker att svordomar
25 dem för att de har hört de äldre göra **det som** lärare, förälder eller annan
26 TV och andra massmedier. MEn **DEt som** jag finner ännu viktigare är
27 bara sitta och just lyssna. Så vad är **det som** får konsekvenser? Frågar jag
28 mot varandra för att sedan ta sällning i **det som** låter klokast och förnuftigast
29 oroa sig för räkningar eller deklaration. **Det som** är så härligt med att vara barn
30 och Gösta Simonsson betecknar **det som** "psykisk svo

Figura 30. Concordancias con *det som* (30 de un total de 48). Corpus de control ARGUS

La figura 31 muestra 15 de las 15 concordancias con la combinación de palabras *den som* en lengua sueca aparecidas en el corpus de control ARGUS. Esta combinación de palabras se corresponde con *el que* en lengua española.

N Concordance

1 markering för att visa ett avstånd mellan **den som** tilltalar och den tilltalade. Jan
2 enbart "vill visa på ett avstånd mellan **den som** tilltalar och den tilltalade."
3 visar ett avstånd mellan två personer. **Den som** tilltalar visar den tilltalade på
4 tanken som räknas. Jag försäkrar er, **den som** snäser av en niande ungdom
5 För övrigt visar erfarenheten att **den som** kan sitt eget språk bra, den har
6 som vill visa på ett avstånd mellan **den som** tilltalar och den tilltalade".
7 säga att språk ska vara ett tillval som **den som** känner ett sug eller behov av att
8 sin far för "gubbhelvete". Jag anser att **den som** använder sig av svordomar och
9 det finns ju en annan sorts slang än **den som** är populär just nu. Den sorten
10 att behandla barn på detta horribla sätt. **Den som** har några invändningar mot mitt
11 och inspirerande livet igenom. **Den som** läser mycket utvecklar ett rikt
12 och kunde läsas om och om igen? **Den som** är uppvuxen med böcker har
13 problem att tänka tillbaka och minnas. **Den som** vuxit upp utan litteratur vet
14 har glädje av att vuxna läser för dem. **Den som** läser får också ökad kontakt
15 mycket så har man lättare att använda **den som** vuxen. En annan sak som är

Figura 31. Concordancias con *den som* (15 de un total de 15). Corpus de control ARGUS

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

La figura 32 muestra 30 de las 275 concordancias con *vad* que se reportan en el corpus de control ARGUS. Esta palabra puede equivaler al español *que, lo que, de que, de lo que* en determinados contextos.

N Concordance

1 ta till mig. Den handlar om att niandet, (*vad* jag förstår oavsett vilken ton man än
2 Han Kritiserar där tilltalsordet "ni" och *vad* han själv uppfattar om niandet och
3 och det är upp till er kära läsare, *vad* ni nu tycker ...
4 synnerhet unga sådana. Guillou vet inte *vad* han talar om, han är fast i någon
5 trots allt kompetentare individer men *vad* gör det om det inte finns några
6 eftersom en arbetslös kostar mer än *vad* arbetsgivaravgiften drar in. Tyvärr
7 så att man kan få mer i socialbidrag än *vad* man tjänar som tex lokalvårdare. Vi
8 Guillou borde läst Olle Josephson text "*vad* skulle Agda Östlund ha gjort med
9 har utelämnat stora delar om *vad* saken egentligen handlar om. Den
10 bilden av elever, men anpassa den till *vad* huvuddelen faktiskt tycker. Gränsen
11 att det är felaktigt att dua folk men *vad* ska vi då tilltala folk i?. Jag
12 ju inte slänga ur mig till famor att: Fan *vad* cool tv du köpt. Om hon nu köpt en
13 om de enda attributerna i svenskan. *Vad* jag menar är inte att svärandet ska
14 exempel blir resonemangen tydliga. *Vad* som saknas är att du inte tar
15 en värdad stockholmsdialekt sagt: Fan *vad* cool tröja, den var faktisk jävligt
16 social fantasi där barn lär sig att tolka *vad* människor känner och tänker, hur
17 lång arbetsdag för du vet antagligen inte *vad* det är. eftersom du har så mycket tid
18 av den tillsägelsen som du kritiserar. *Vad* ger dig rätten att klaga på den här
19 kanske var det det," du syftade här på *vad* pojken hade gjort för att förtjäna
20 vet man in om det viktigaste är *vad* jag håller på med eller om det
21 vill lyfta fram i sitt berättande. Ex.: "*Vad* faan håller du på med din jävla
22 jobbigt att lyssna på, man vet då inte *vad* personen verkligen vill lyfta fram i sitt
23 diskuteras av diverse "förstå-sig-påare". *Vad* dessa människor har gjort, är att
24 svara på den frågan måste man reda ut *vad* "fula ord" är. Enligt min mening är
25 stämmer enligt min mening bra. *Vad* jag skulle vilja tillägga är att även
26 Detta just pg.a. sitt språkbruk. Är detta *vad* en kurs i retorik går ut på? Är detta
27 sätt än med fula ord och svordomar. *Vad* kan vi då göra åt problemet? Att helt
28 Ett lätt och snabbt sätt att ta reda på *vad* som händer ute i världen. Andra
29 kanske har rätt, Ame Rehberg. Men *vad* ska man då göra är det? Att försöka
30 är nog ett omöjligt uppdrag, Men *vad* vi kan göra är att ta ansvar för våra

Figura 32. Concordancias con *vad*. Corpus de control ARGUS

RESULTADOS CUANTITATIVOS DE LOS CORPUS

La figura 33 muestra 20 de las 46 concordancias aparecidas en el Corpus de control ARGUS para la combinación *vad som*, equivale a las combinaciones *lo que*, *el cual*, *el que* en español.

N Concordance

1	liv och hjälper dem att själva förklara <i>vad som</i> känslomässigt pågår i deras liv,
2	mindre tid som barn ägnar åt läsning, <i>vad som</i> kan göras åt det och hur det
3	förälder pratar mycket med barnen om <i>vad som</i> har hänt under dagen osv. Att
4	ville prata med mina föräldrar och veta <i>vad som</i> var "rätt och fel". Det slutade
5	bara till härliga minnen, de lär även barn <i>vad som</i> är rätt och fel, de förklarar
6	många som säger att den är bra. Det är <i>vad som</i> hände mig när jag gick i
7	hjälp till att ge en förståelse om <i>vad som</i> är rätt och fel. Självt, fick jag
8	hur han hatade att läsa som barn och <i>vad som</i> blev dess konsekvenser. —
9	men för att utveckla tankar, känslor, <i>vad som</i> är rätt eller fel, för att göra oss
10	hjälp av böckerna vad fantasi är. Och <i>vad som</i> kan föras från fantasi till
11	min synvinkel försöka reda ut vem eller <i>vad som</i> är boven i dramat. Astrid
12	verklighets uppfattning och lär dem <i>vad som</i> är rätt och orätt. Genom att
13	trösta och ge dem sina synpunkter på <i>vad som</i> är rätt och fel. På så sätt lär sig
14	oss med, om vi inte begriper och förstår <i>vad som</i> skrivs i tidningar, böcker? Då
15	De yngre som inte har någon talan om <i>vad som</i> ska ses på tv får också ta del
16	där bara fantasin sätter gränser för <i>vad som</i> kan hända. Vi får lära oss om
17	"Herr Nilsson" och hon gör i princip <i>vad som</i> faller henne in, kort sagt så är
18	Genom att läsa får barnet veta <i>vad som</i> är rätt och fel, sant och falskt.
19	utöka ordförrådet om känslor, regler för <i>vad som</i> är rätt och fel och om
20	han satte sig tom ner och berättade <i>vad som</i> hände med Elisabeth och

Figura 33. Concordancias con *vad som/lo que/el cual/el que*. Corpus de control ARGUS

La figura 34 muestra 35 de un total de 74 agrupaciones de palabras/*clusters* con *som*, aparecidas en el corpus de control ARGUS.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

N	Cluster
1	VAD SOM äR
2	BARN SOM ARBETAR
3	är NåGOT SOM
4	BARN SOM HAR
5	DE BARN SOM
6	Går I SKOLAN
7	SOM INTE Går
8	är DET SOM
9	SPRåk SOM GER
10	ETT BARN SOM
11	DET SOM är
12	SOM HOTAR DERAS
13	BARN SOM INTE
14	INTE Går I
15	ETT SPRåk SOM
16	OCH VAD SOM
17	MöNSTER FÖR FORTSATT
18	HOTAR DERAS HälSA
19	LYSSNA På SPRåk
20	SOM MAN KAN
21	SOM JAG INTE
22	SOM T EX
23	SOM PROSTITUERAR SIG
24	SOM BARN OCH
25	På SPRåk SOM
26	SOM är Rätt
27	SOM GER DEM
28	EN BOK SOM
29	ATT LYSSNA På
30	DERAS HälSA UTBILDNING
31	Rätt OCH FEL
32	SOM JAG TYCKER
33	SOM FINNS I
34	DEM MöNSTER FÖR
35	VAD är DET

Figura 34. Agrupaciones de palabras/*clusters* con *som*. Corpus de control ARGUS

La figura 35 muestra 30 de las 50 líneas de concordancia de: *vilket/que/lo que*.

RESULTADOS CUANTITATIVOS DE LOS CORPUS

N Concordance

1 tid för att laga bil & hus på egen hand, *vilket* skulle bli billigare. Mindre arbetstid
 2 uppfattar niandet, som ett sätt att visa *vilket* status eller klass man tillhör.
 3 Sie, som betyder ni på svenska. *Vilket* språk, förutom engelskan, har
 4 inte att de skadar varandra på riktigt, *vilket* dom heller inte gör. Det är alltså
 5 ihop 40 000 namn underskrifter, *vilket* är ganska mycket. men knappast
 6 eftersom det bör gå lika bra att skriva *vilket* ord som helst med full kraft. När
 7 Så som TVs lätt användliga interface. *Vilket* barn som hållst vet hur man sätter
 8 Ett medföljande problem är utbildning *vilket* många aldrig får. I Barnen och vi
 9 Åldern varierar från 5 år och uppåt, *vilket* betyder att en del har 10 år tills de
 10 mening skulle kunna vara "Fy fan, *vilket* fint väder idag!" Inget konstigt me
 11 måste man vara försiktigt beroende på *vilket* sällskap man befinner sig i, och
 12 inte alls vara jordens undergång *vilket* herr Mullhauser tycks antyda.
 13 1989 Antas Barnkonventionen, *vilket* innebär att barn har "rätt att
 14 hushåll arbetar 10—15 timmar per dag, *vilket* är mer än en vuxen i ett i-land. Det
 15 sina pengar på annat sätt, t.ex stjäla. *Vilket* är bäst? Att barnen arbetar eller
 16 hemma gör. 5 % av alla arbetande barn *vilket* är 10 miljoner arbetar inom
 17 skulle det kosta 48 miljarder kr per år *vilket* motsvara mindre än 1 % av
 18 arbetsskydd och för kort dygnsvila *vilket* leder till många dödsfall. Det mest
 19 "NU JÄVLAR ÄR JAG FÖRBANNAD" *vilket* tror ni att man reagerar mest på?
 20 år arbetar. Detta är förbjudet enligt ILO, *vilket* betyder Internationella
 21 Barnen har omänskliga arbetstider *vilket* medför dålig dygnsvila och då blir
 22 verkar vara helt emot svordomar, *vilket* i och för sig är bra. Det är alltid bra
 23 då bli ensamma med sina egna tankar, *vilket* vore förödande för många. Nej, jag
 24 Ibland kan de ha en språklig funktion, *vilket* de har i meningen "Du är
 25 *vilket* leder till mindre tid för barnen, *vilket* kan leda till större konsekvenser i
 26 tidigare så är dagens jobb stressigare *vilket* leder till mindre tid för barnen, *vilket*
 27 gick bort när hon var mycket liten, *vilket* även min pappa gjorde. Författare
 28 språkets grammatiska strukturer" *vilket* Kerstin Rimsten-Nilsson påpekar i
 29 som påverkas utav att läsa bamböcker, *vilket* jag tycker betonar vikten av att
 30 har höga positioner på arbetsplatser *vilket* leder till att arbetet tar mycket tid

Figura 35. Concordancias con *vilket* (30 de un total de 50). Corpus de control ARGUS

La figura 36 muestra el total de las 29 líneas de concordancias con *vilken/que/lo que*.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

N Concordance

1 om att niandet, (vad jag förstår oavsett *vilken* ton man än använder) låter

2 om man niar någon, så visar man från *vilken* "högre" klass man kommer ifrån.

3 alla behandlas olika, beroende *vilken* familj eller klass man tillhörde.

4 födda efter 1960, är ovetande om *vilken* nedlåtande klang niandet har, utan

5 Vi kan gemensamt skapa den värld i *vilken* alla barn har lika rättigheter,

6 i en artikel ("Att få välja sina ord — *vilken* lyx" Svd 22.2.1997). "Språklig

7 Om alla dessa ord skulle vara fula, på *vilken* skala ligger då alla dessa svårord?

8 alfabetets hav. På ön står en vulkan, ur *vilken* all världens hemligheter sprutar.

9 tydligt din åsikt: Jag vet inte det. *Vilken* isåfall? Ge argument: Här finns

10 den sneda bilden på barn i U-länder. *Vilken* är då denna faktor. (paus) Jag

11 i Latinamerika. Man kan fråga sig själv, *vilken* framtid kommer dessa barn att få?

12 är "När används detta uttryck?" och "*Vilken* funktion har detta uttryck i

13 sittande regeringen, jag minns tyvärr ej *vilken*, att privata radiostationer skulle få

14 ut till en bred och medelålders publik. *Vilken* typ av radio resulterar detta i? Ja,

15 istället att man borde fundera över *vilken* roll de spelar som uttrycksmedel.

16 för att fördömma ska ställa sig frågan *vilken* funktion sådana här ord fyller i

17 vad man är för slags människa och *vilken* grupp man tillhör. Jag håller med

18 använda svordomar och fult språk och *vilken* betydelse har det för att vi ska

19 om vad man är för människa och till *vilken* grupp man tillhör genom

20 ELLER MINDRE GLAD BERODDE PÅ *VILKEN* BOK VI HADE FÅTT HÖRA.

21 kunde frigöra mig från mina föräldrar. *Vilken* triumf det var när jag själv

22 För bamboken! *Vilken* var eran favorit stund på dagis?

23 läsa böcker, vuxna som barn. Tänk på *vilken* uppsjö av fina författare vi har i

24 och läslust, 1981 förstår man *vilken* otrolig möjlighet vi har att hjälpa

25 läste en artikel som visar på indirekt *vilken* nytta rappartisten Petter skulle

26 fall att det skulle inte spela någon roll *vilken* bok jag valde. Bara jag läste

27 liten? Så minns jag den tid som var, *vilken* nu känns som en fjärran värld.

28 Oavsett varifrån man kommer, *vilken* kultur man har, kan och bör man

29 att jag sett Pippi ledsen, oavsett *vilken* situation hon befann sig i, kan ni?

Figura 36. Concordancias con *vilken* (29 de 29). Corpus de control ARGUS

La figura 37 muestra el total de las 18 concordancias con *vilka/los que*.

N Concordance

1 svar, publicerade i Språkvård nr3 1997, *vilka* behandlade svenska språket som

2 essäer, dikter eller grammatikböcker av *vilka* kurserna i svenska idag består.

3 studerande känner jag många elever *vilka* har stora svårigheter med rent

4 tänka efter lite så kanske man märker *vilka* tillfälle man kan använda ordet ni i.

5 betyder inte att man kan använda *vilka* ord som helst. Att använda könsord

6 på t.ex Nobel middagen ni kan tänka er *vilka* reaktioner det skulle bli på

7 är? Hur vi ska använda de olika och *vilka* kan vi lita på? Massmedia finns runt

8 vi har vidtagit är att bojkotta de varor *vilka* är tillverkade av företag som

9 Men det kommer alltid att finnas, *vilka* lagar vi än använder. Jag hoppas

10 *vilka* ställen män vill bli klämda på och *vilka* ställen kvinnor vill bli klämda på. Är

11 utan att det finns artiklar om *vilka* ställen män vill bli klämda på och

12 varför dessa händelser sker och *vilka* följderna blir. Detta gör att barnen

13 Så småningom lär dem sig även i *vilka* sammanhang orden används.

14 ord lärs in. Ord ger barn frågeställningar *vilka* aktiverar deras tankar och fantasi,

15 våldet i samhället till uppväxten och *vilka* förebilder man har. Även språket

16 privat under hela livet. Visa eleverna *vilka* olika typer av böcker det finns så

17 yngre, äldre kan sälla bort våld och inse *vilka* konsekvenser spelandet kontra

18 och det kan vara svårt att veta *vilka* det gäller av det stora utbudet.

Figura 37. Concordancias con *vilka* (18 de 18). Corpus de control ARGUS

RESULTADOS CUANTITATIVOS DE LOS CORPUS

La figura 38 muestra la totalidad de las 6 concordancias aparecidas para la combinación *säger att/dice que*, una de las combinaciones idiomáticas objeto del presente estudio (la menos frecuente en el corpus de control ARGUS).

- N Concordance
- 1 som möjligt? Vissa går nog i taket, och säger att kunskap i det svenska språket
 - 2 svära. Det blir inte samma sak om du säger att tröjan var jättefin. Hur mjäkigt
 - 3 lagar som förbjuder barnarbete och säger att alla barn har rätt till en
 - 4 Det är det jag menar när jag säger att svordomar är bra att ha i det
 - 5 flesta om, men de bara igonera det och säger att det inte ha med dem att göra,
 - 6 klassiker för att det är så många som säger att den är bra. Det är vad som

Figura 38. Concordancias con *säger att*. Corpus de control ARGUS

El Diagrama 3 muestra el resultado de la comparación entre el corpus SAELE y el corpus CEDEL2 en cuanto a: cantidad total de palabras; concordancias con *que, lo que, de que, algo que* y *dice que*. Estos resultados cuantitativos se discuten sucesivamente en el capítulo 7.

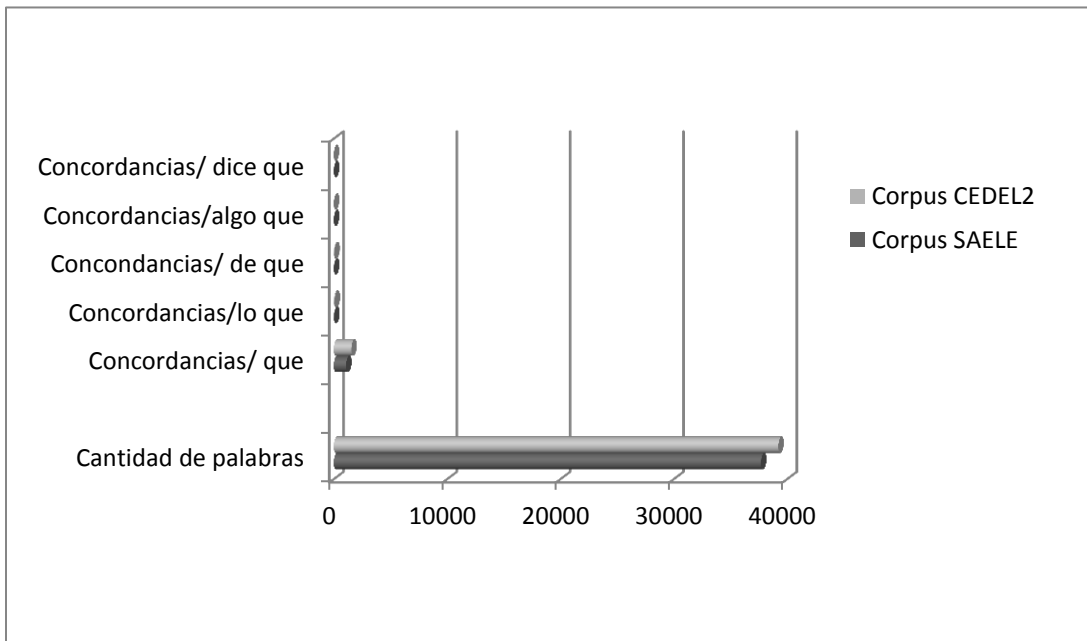


Diagrama 3. *Que* y combinaciones de palabras con *que*. Corpus SAELE y CEDEL2

El diagrama 4 presenta la correlación existente entre todos aquellos elementos lingüísticos del corpus de control ARGUS que en uno u otro contexto fraseológico equivalen a *que* o a las combinaciones de palabras con *que*: *lo que*, *de que*, *algo que* y *dice que*, objetos de análisis en la presente investigación. Como se aprecia *att* (2908) ocupa el primer lugar en frecuencia seguido de *som* (1652); *vad* (275); *vilket* (50); *det som* (48); *vad som* (46); *något som* (33); *vilken* (29); *vilka* (18); *den som* (15); *någon som* (14); y *säger att* (6).

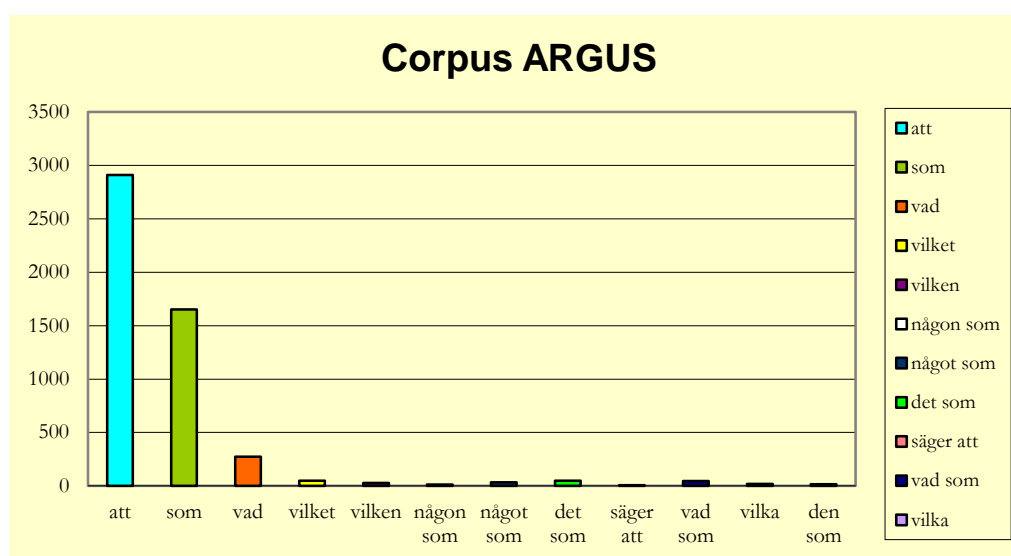


Diagrama 4. Frecuencia de palabras equivalentes a *que/att*, *som*, *vad*, *vilket*, *vilken*, *vilka*; *lo que/vad*, *vad som*, *någon som*, *något som*; *de que/det som*, *den som*, *algo que/något som*, *någon som* y *dice que/säger att*. Corpus de control ARGUS

6.4 Resultados de la comparación de frecuencias entre el corpus SAELE y el subcorpus de nativos hispanohablantes del CEDEL2

Los principales resultados de la aplicación del calculador Log-likelihood aparecen en los cuadros a continuación. El Cuadro 16 muestra que la frecuencia de aparición de *que* recopiladas en el CEDEL2 (subcorpus de hispanohablantes) es superior a la aparecida en el Corpus SAELE.

RESULTADOS CUANTITATIVOS DE LOS CORPUS

Cuadro 16. SAELE versus CEDEL2. Frecuencias *que*

Key:

O1 is observed frequency in Corpus 1 (Corpus SAELE)- *que*
 O2 is observed frequency in Corpus 2 (Corpus CEDEL2)- *que*
 %1 and %2 values show relative frequencies in the texts.
+ indicates overuse in O1 relative to O2,
- indicates underuse in O1 relative to O2

Item	O1	%1	O2	%2	LL
Word	1094	2.95	1550	3.95 -	55.33

La combinación *lo que* aparece como la más frecuentemente empleada por nativos hispanohablantes en el subcorpus de control CEDEL2, como se aprecia en el Cuadro 17, cuando se calcula comparativamente las frecuencias de las combinaciones de palabras con *que*: *lo que*, *de que*, *algo que* y *dice que*. La aparición de la misma en el corpus SAELE muestra que los aprendices suecos de E/LE utilizan la misma en una frecuencia equivalente a la mitad comparado con los hispanohablantes.

Cuadro 17. SAELE versus CEDEL2. Frecuencias *lo que*

Key:

O1 is observed frequency in Corpus 1SAELE
 O2 is observed frequency in Corpus 2CEDEL2
 %1 and %2 values show relative frequencies in the texts.
+ indicates overuse in O1 relative to O2,
- indicates underuse in O1 relative to O2

Item	O1	%1	O2	%2	LL
Word	40	0.11	82	0.21 -	12.48

En situación aún más desventajosa según el cuadro 18, aparece el uso de la combinación *de que*, utilizada entre los aprendices suecos de E/LE del corpus SAELE en un 0.05%, mientras que los nativos hispanohablantes del CEDEL2 la emplearon en un 0.16%.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

Cuadro 18. SAELE versus CEDEL2. Frecuencias *de que*

Key:

O1 is observed frequency in Corpus 1SAELE

O2 is observed frequency in Corpus 2CEDEL2

%1 and %2 values show relative frequencies in the texts.

+ indicates overuse in O1 relative to O2,

- indicates underuse in O1 relative to O2

Item	O1	%1	O2	%2	LL
Word	17	0.05	62	0.16 -	24.75

Resumiendo: el menor uso de las combinaciones *lo que* y *de que* por los aprendices suecos de E/LE del corpus SAELE contrasta con el sobreuso de combinaciones tales como *algo que* (Cuadro 19) y *dice que* (Cuadro 20), respecto al uso de los nativos hispanohablantes del CEDEL2.

Cuadro 19. SAELE versus CEDEL2. Frecuencias *algo que*

Key:

O1 is observed frequency in Corpus 1SAELE

O2 is observed frequency in Corpus 2CEDEL2

%1 and %2 values show relative frequencies in the texts.

+ indicates overuse in O1 relative to O2,

- indicates underuse in O1 relative to O2

Item	O1	%1	O2	%2	LL
Word	13	0.04	6	0.02 +	3.05

RESULTADOS CUANTITATIVOS DE LOS CORPUS

Cuadro 20. SAELE versus CEDEL2. Frecuencias *dice que*

Key:

O1 is observed frequency in Corpus 1SAELE

O2 is observed frequency in Corpus 2CEDEL2

%1 and %2 values show relative frequencies in the texts.

+ indicates overuse in O1 relative to O2,

- indicates underuse in O1 relative to O2

Item	O1	%1	O2	%2	LL
Word	15	0.04	0	0.00 +	21.66

Hasta aquí hemos mostrado los principales resultados cuantitativos del corpus SAELE y los corpus de control: subcorpus ARGUS (nativos suecos) y el subcorpus CEDEL2 (nativos hispanohablantes). Los datos obtenidos con la ayuda de las herramientas computacionales los analizaremos más profundamente en el capítulo de discusión (Cap. 7).

CAPÍTULO 7. DISCUSIÓN DE LOS RESULTADOS

A partir de los objetivos del presente estudio, el análisis de los resultados cuantitativos obtenidos en el corpus de aprendices suecos de E/LE (SAELE) en contraste con los corpus de control CEDEL2 (subcorpus de nativos hispanohablantes) y ARGUS (subcorpus de nativos suecos) y basados en los soportes metodológicos de la LC y el ACI, se discuten en el presente capítulo el uso y la función de las combinaciones de palabras con *que: lo que, de que, algo que y dice que* y sus equivalentes suecos para realizar un ACI del aprendiz sueco de E/LE.

Primeramente analizamos el uso de las combinaciones de palabras con *que (lo que, de que, algo y dice que)* por los aprendices suecos de E/LE (Corpus SAELE), ya que, según los objetivos, es ésta la pregunta central del presente estudio (objetivo 1). Seguidamente se analiza la frecuencia y uso de dichas combinaciones de palabras en los corpus de control CEDEL2 (subcorpus de nativos hispanohablantes) y ARGUS (subcorpus de nativos suecos) contrastando los mismos con los resultados obtenidos en el Corpus SAELE (objetivos 2 y 3). El propósito principal es mostrar las diferencias en tipos y frecuencias de usos que se presentan entre nativos hispanohablantes y suecos en relación con las producciones de nativos suecos en camino hacia la lengua meta, el E/LE. El resultado del análisis se emplea posteriormente para discutir eventuales influencias de la LM de los aprendices en la producción de textos en E/LE. Sucesivamente se aborda la información extraída de la encuesta efectuada que nos ofrece los datos sociolingüísticos de los participantes (objetivo 4), y ya teniendo todos estos elementos valoramos si coincide la descripción del nivel de destreza del aprendiz (A2-B1) con su competencia real (objetivo 5), para finalmente terminar con la descripción del perfil del aprendiz sueco de E/LE de nuestro estudio.

7.1 Uso de las combinaciones de palabras con *que: lo que, de que, algo que y dice que* por los aprendices suecos de E/LE

El objetivo general del estudio ha sido comparar el uso de algunas combinaciones de palabras de la producción escrita de aprendices adultos suecos de E/LE (nivel A2-B1

DISCUSIÓN DE LOS RESULTADOS

del MCERL) con la de nativos hispanohablantes y suecos. Específicamente, describir algunas características de la producción escrita de aprendices adultos suecos de E/LE a través de identificar la frecuencia y el uso de combinaciones de palabras con *que: lo que, de que, algo que, dice que* en textos argumentativos recopilados en dos universidades suecas, durante los cursos lectivos 2008/09 y 2009/10 que forman parte del corpus de aprendices SAELE, según planteamos en el primer capítulo.

Para dar respuesta a este primer objetivo nos centramos en cómo utilizan los aprendices suecos de E/LE las combinaciones de palabras con *que: lo que, de que, algo que y dice que* compiladas en el corpus SAELE⁹¹.

Según aparece en el capítulo de resultados (6)⁹² el corpus SAELE tiene un total de 37.076 palabras, donde la palabra *que* ocupa el quinto lugar con una frecuencia de 1.094 (2.95%). No se apreciaron diferencias significativas entre los dos minicorpus (uno para cada universidad) en relación con el número total de textos en que aparece la palabra *que* y sus combinaciones, en sus formas esenciales de usos ni en la frecuencia de su aparición. No obstante se debe señalar que en el corpus SAELE hay recopilados 89 textos de aprendices de la universidad de Gotemburgo y 46 del instituto superior de Jönköping⁹³ (cuadro 9).

En el Corpus SAELE aparecen ejemplos que ratifican que el uso de *que* por los aprendices suecos de E/LE (corpus SAELE) coincide con las funciones descritas acerca de su valor de conjunción, pronombre relativo, interrogativo, exclamativo, y comparativo, así como de que al unirse a otras palabras puede formar locuciones conjuntivas.

Por una parte, *que* en su función conjuntiva, se une a otras estructuras lingüísticas como conjunciones, preposiciones y adverbios, etc. y, por otra, en su papel de relativo denexo subordinante, también cumple una u otra función en contexto a partir de la relación que establece con las palabras/estructuras que le rodean e incluso puede que no quede claro su papel en una función o en otra, tal y como indica Alarcos (Alarcos, 1994: 260).

Que puede equivaler en sueco a: *att, som, vilken/vilket/vilka, vad y än*, partículas que cumplen funciones similares al *que* del español como aparece en el cuadro 3 (Cap.4).

⁹¹ La explicación acerca del corpus SAELE aparece en el capítulo 5.

⁹² Véanse los cuadros y diagramas.

⁹³ La información aparece recogida en la base de datos del corpus SAELE.

Si se considera la función múltiple de la partícula *que* estamos de acuerdo en que sería imposible abarcar en esta discusión todos sus valores por las limitaciones de espacio y objetivos del trabajo, pero sí destacaremos lo más relevante en el uso de las diferentes combinaciones de palabras con *que* por aprendices suecos de E/LE, que en la práctica ha sido mucho más variado y amplio de lo que en realidad esperábamos debido a que los aprendices utilizan combinaciones de palabras que coinciden con el principio idiomático o fraseológico, ‘crean’ y muestran ‘un uso de la lengua’ propio del *continuum* interlingüístico en que se encuentran (cf. Sinclair, 1996; Cook, 2002).

En la lista de concordancias de las combinaciones de palabras con *que* (Cap. 6, cuadro 10) se observa que la combinación *lo que* aparece más frecuentemente (n=40), seguida de *de que* (n=17), *dice que* (n=15), siendo la menos utilizada *algo que* (n=13). Cabe destacar que este resultado coincide con los dos estudios anteriores donde se reportó la combinación *lo que* como la de mayor frecuencia de uso tanto entre aprendices suecos de E/LE que se encontraban en un nivel B1-B2 (cf. Pino, 2006) como entre los que se encontraban en un nivel de aprendizaje similar a los participantes de este estudio (cf. Pino, 2010). Otro elemento interesante en la comparación entre este estudio y los anteriores reside en que los aprendices, cuyo nivel se sitúa entre el A2-B1 utilizaron las combinaciones *de que*, *algo que* y *dice que* en proporción similar, mientras que los aprendices situados en un nivel más avanzado las utilizaron en proporción menor a estos.

La cantidad de veces que la palabra *que* se emplea en los 134 textos de los 135 que posee el corpus SAELE muestra la variación cuantitativa en el uso de *que* por los aprendices. Esto se refleja en la cantidad de veces que esta palabra aparece en los diferentes textos (Figura 10), desde la utilización de un solo *que* (Figura 12) hasta la aparición de la palabra un total de 27 veces (Figura 11) y cualitativa respecto a cómo utilizan la partícula en sus funciones como conjunción, pronombre y comparativo.

Es factible afirmar que se observa una relación directa entre el número total de palabras en los textos y la cantidad de veces que la palabra *que* y sus combinaciones aparecen y ello indica probablemente la tendencia del uso de dicha palabra y sus combinaciones en relación proporcional con el mejor dominio del idioma y su vocabulario. Si por ejemplo, comparamos los datos de colocaciones con *que* entre el corpus SAELE y el CEDEL2 (subcorpus de nativos hispanohablantes) hay diferencias entre colocación/ número de texto en que aparecen. Por ejemplo: el

DISCUSIÓN DE LOS RESULTADOS

corpus SAELE muestra una correlación de 103/135 textos/colocación *de que* mientras el de nativos hispanohablantes muestra 125/135/colocación *de que*; por otra parte observamos una correlación de 80/51 textos/colocación *lo que* (corpus SAELE) y 176/73 textos en el subcorpus de nativos hispanohablantes del CEDEL2 (Figuras 7 y 12).

La palabra *que* y sus combinaciones aparecen utilizadas correctamente en varias de sus funciones, en diferentes textos del corpus SAELE, como en el caso de uso de *que* comparativo en el texto del cuadro 13:

(53) México tiene un coeficiente de Gini mucho más alto que el resto de los miembros de OCDE

En otra parte del texto el aprendiz olvida el cuantificador (mayor, menor e igual)

(54) México puede obtener un nivel de desigualdad *que* el resto de la OCDE.

De modo más específico en la comparación entre el texto (cuadro 11) y el texto (cuadro 12) cabe destacar que el primero está formado de oraciones compuestas largas y el aprendiz utiliza *que* como conjunción y relativo y conectores como *pero* y *aunque* para enlazar las frases, mientras que en el segundo (Cuadro 12), *que* se utiliza en una oración relativa con función explicativa en una expresión corta y concisa en la cual aparece además la conjunción *y* normalmente para enlazar las frases excepto cuando utiliza *que* en la oración relativa donde además se vale de *pero*. Esta diferencia que se aprecia entre los dos textos en cuanto a cantidad y usos de *que* denotan en nuestra opinión el nivel individual de aprendizaje en que se encuentra el aprendiz, a pesar de los errores que se cometen, algunos de ellos en relación con *que* en contexto con las palabras adyacentes tal como mostraremos a continuación.

(55) Entre los miembros de la OCDE México está en el primer puesto en la lista a pesar de que el nivel de pobreza y desigualdad ha caído durante los últimos años

En la relación contextual aparece *que* con el subordinador concesivo *a pesar de* sucedido de un verbo en indicativo cuando debía estar en modo subjuntivo *haya caído* teniendo en cuenta que la frase refiere primero un hecho concreto, el cual el escribiente refuta a continuación. En sueco sería *faktum och avståndstagande*. La elección del subjuntivo en este caso está determinada porque el hablante se refiere a un 'hecho cierto' y seguramente piensa que el lector ya conoce ésta. Se emplearía el indicativo si el hablante sabe que es verdad lo que dice y quiere informar al oyente porque considera que este desconoce la información (cf. Martí & Torrens, 2001). El

error asociado al uso del modo subjuntivo es usual en los aprendices suecos de E/LE debido a que prácticamente no utilizan este modo en su lengua materna (cf. Söhrman, 2007).

En otros dos ejemplo tomados del texto del cuadro 14 nos podríamos encontrar con un error similar, aunque en el ejemplo 56 podría emplearse también el pretérito imperfecto del verbo haber en el modo indicativo.

(56) A pesar de *que* ella tiene un tiempo duro, siempre *está en* buen humor y *es* muy conciderado

En lengua española diríamos ‘A pesar de que ella hubiera pasado por /había pasado por un tiempo/una época difícil siempre estaba de buen humor y era muy considerada’. Si analizamos el texto como conjunto vemos que el aprendiz para contar sobre la película utiliza el presente de indicativo: trabaja, tiene, sabe etc., pero en la medida que la narración transcurre llega a un punto en que hace uso de *a pesar de que* para refutar un hecho por lo que en este caso también podría usar el modo subjuntivo. La posibilidad de utilizar el modo indicativo o el subjuntivo cuando estamos en presencia de *a pesar de + que* causar dificultades a los aprendices suecos de E/LE participantes en este estudio.

(57) Bill confiesa y pide a Selma *que* no va a decir-- a la mujer de él, Linda/Bill bekänner och ber Selma att hon inte berättar det för hans hustru, Linda

En este ejemplo el aprendiz escribe: ‘*que* no va a decir a la mujer de él, Linda’ en lugar de: ‘*que* no vaya a decirle a Linda su mujer/la mujer de él’ y en este caso utiliza *va*, la tercera persona del presente de indicativo del verbo ir, en lugar de *vaya*, tercera persona del presente del modo subjuntivo.

Los errores en el uso del subjuntivo caracterizan la interlengua del aprendiz sueco en este nivel de aprendizaje y como explicamos en el capítulo 4 esto se debe a que el modo subjuntivo prácticamente no se utiliza en lengua sueca, y el hecho de utilizar otros recursos que dominan mejor (modo indicativo) por su relación más cercana a su LM les permite en ocasiones construir o ‘reconstruir’ combinaciones y frases como ‘el hecho de tener un tiempo duro’ que para un nativo no sería la más natural como se reporta en estudios anteriores (cf. Wray, 2005).

Encontramos otros ejemplos seleccionados (*concordance* con *que*) donde se observa la partícula *que* empleada correctamente respecto al modo verbal indicativo. Por ejemplo:

DISCUSIÓN DE LOS RESULTADOS

(58)

N Concordance

152 Quiero aprender del país y la ciudad en **que** estoy y tambien quiero vivir la

N Concordance

34 Había mucha pobreza y muchos niños **que** trabajaban por las calles y esto nos

N Concordance

626 Yo no estoy seguro, pero creo **que** ella es una prostituta. Si yo había

En la línea 152 (presente de indicativo de estar); en la línea 34 (pretérito imperfecto de indicativo de trabajar) y en la 626 (presente de indicativo de ser).

Las 34 agrupaciones de tres palabras extraídas en el corpus SAELE con la herramienta *clusters* (Figura 9, Cap. 6) nos muestran que hay una interrelación entre estas combinaciones de palabras y los temas de las composiciones sobre los que han escrito los aprendices suecos de E/LE. Obsérvese por ejemplo la combinación *que se llama* (n=30), *que me gusta* (n=25), *creo que es* (n=12), *me parece que* (n=10), y *creo que* (n=7) donde en la primera hay un verbo de identificación y las cuatro últimas hay verbos de opinión que anteceden a *que*. Todas ellas tienen en común que se utilizan para expresar la opinión sobre un hecho que necesita ser contado y explicado, ya sea este ‘Mi peli favorita’, ‘Mi tiempo libre’ o ‘El deporte que prefiero practicar’, ‘Vacaciones inolvidables’, ‘Recuerdos de Navidad’ o la argumentación de un artículo periodístico, que son los temas elegidos para escribir los textos de forma argumentativa (véase el Cap. 5).

Otro aspecto que merece atención en el corpus SAELE son las llamadas *collocates* (Fig.7, cap. 6). Aquí vemos que las más utilizadas a la derecha de *que* por los aprendices son: *que se* (n=74), *que el* (n=44), *que es* (n=43) y a la izquierda tenemos a: *lo que* (n=40), *creo que* (n=38) y *es que* (n=24). También aparecen *de que* (n=17), *dice que* (n=15) y *algo que* (n=13). Al continuar indagando nos percatamos que de un total de 74 líneas con *que se* ésta aparece combinada 32 veces con el verbo *llamar* y esto indica también una relación con los textos que escriben los aprendices donde explican por ejemplo cómo se llama un lugar o se refieren al nombre del protagonista de una película. Por ejemplo: [En Málaga hay un teatro antiguo, el teatro romano *que se llama* La Alcazaba] y en [Al principio un hombre, *que se llama* Forrest Gump] (corpus SAELE).

Seguidamente vamos a comparar los textos producidos por los aprendices suecos con los producidos por nativos hispanohablantes del subcorpus del CEDEL2 y el subcorpus del ARGUS respecto al uso de las combinaciones de palabras con *que*.

7.1.1 Uso de la combinación *lo que* por los aprendices suecos de E/LE

El análisis de las 40 concordancias de *lo que/vilken/vilket/något som/vad som/det som/så*, en el corpus SAELE (Figura 3, Cap. 6) muestra que el tipo de palabra que más frecuentemente precede a esta combinación son verbos (n=11), sustantivos (n=6), preposiciones (n=4), conjunciones (n=5), adverbios (4), conectores (3) y adjetivos. (n=3). De los verbos *saber* (n=3) es el que más se utiliza. Asimismo se registra que la mayoría de las palabras que suceden a la combinación *lo que* son verbos (n= 36) y el resto son artículos (n=1), adverbio (n=1), nombre (n=1) y sustantivo (n=1). Entre las palabras que preceden a *lo que* que forman una combinación recurrente están *pero* (encabezando la oración) + *lo que* (n=2), *y* + *lo que* (n=2) y *todo* + *lo que* (n=2) y el verbo que predomina es: *saber* (3); entre las que la suceden las más llamativas son los verbos *representar* (n=2), *experimentar* (n=1) y *sentir* (n=1) (verbos de emoción) y *gustar* (n=2) (verbos de opinión).

Los nativos hispanohablantes (subcorpus de CEDEL2) utilizan en primer lugar verbos (n=15) como precedentes de *lo que* y en segundo lugar se destaca la combinación *por* + *lo que* (n=11) como se aprecia en la figura 17. Además, aparecen las preposiciones *de* (n=8), *en* y *con* (n=4), y la conjunción *y* (n=4). Algo distintivo es que *lo que* aparece introduciendo la oración 12 veces. Entre los verbos que preceden a *lo que* están *saber* (n=1), *hacer* (n=2) y *ser* (3) y le suceden *soñar* (n=1), *querer* (n=4), *desear* (n=1) y *llorar* (n=1) (verbos de emoción) y *gustar* (n=1), *opinar* (n=1), (verbos de opinión).

Por su parte, la mitad de los nativos suecos del corpus ARGUS prefieren la combinación *är/ser* + *algo que* cuando *algo que* equivale a *något som* (Figura 28) y utilizan *någon som/algo que* sólo esporádicamente (Figura 29). Entre los verbos que suceden a *något som* se destaca *tycka/opinar* (n=3) (verbo de opinión) y *oroa/preocupar* (n=1), *drabba/afectar* (n= 1), *förbättra/mejorar* (n=1), *reta/ofender* (n=1), *beundra/admirar* (n=1) (verbos relacionados con un estado mental positivo o negativo).

Resumiendo este aspecto podemos plantear que se aprecia una tendencia que muestra que los aprendices suecos de E/LE (corpus SAELE) coinciden con los nativos hispanohablantes (subcorpus CEDEL2) y los nativos suecos (subcorpus

DISCUSIÓN DE LOS RESULTADOS

ARGUS) en utilizar verbos de emoción y de opinión sucediendo a *lo que*. Por otra parte apreciamos también que los aprendices suecos de E/LE (corpus SAELE) se diferencian de los nativos hispanohablantes en que solamente utilizan 1 vez como antecedente de *lo que* a la preposición *por*, mientras que los nativos hispanohablantes la utilizan 11 veces. Esto indica en buena medida la diferencia entre los textos de nativos hispanohablantes y aprendices si tenemos en cuenta el valor que adquiere *por* en la argumentación junto a otras combinaciones introduciendo la causa de un hecho o una consecuencia (cf. Martí Sánchez, 2008; Martí & Torrens, 2001)

En las concordancias con *lo que*, un aprendiz ha utilizado la expresión idiomática ‘cueste lo que cueste’ (Figura 3, línea 8, Cap. 6), lo que demuestra progreso en la LE, aunque contrasta con las tres empleadas por los hispanohablantes: ‘dar por hecho lo que no está hecho’, ‘ver más allá de lo que pueden ver’ y ‘hacer lo que quiera’ (Cap. 4).

A continuación aparecen las líneas 23, 25 y 28⁹⁴ de la concordancia con *lo que* del corpus SAELE.

N Concordance

23 como Mandela, es el unica gurdia **lo que** entiende las cartas que Mandela

N Concordance

25 nos encontramos a un chico holandés **lo que** compartió taxi con nosotros. Está

N Concordance

28 chóferes”. El tercer pensamiento es **lo que** pasará en las comisarías cuando

Tal como se aprecia en los ejemplos hay errores en la elección de la combinación *lo que* elegida por parte de los aprendices suecos de E/LE ya que en estos contextos deberían haber seleccionado: *que* (línea 23) porque hay un antecedente expreso ‘el único guardia’; *que/el que/el cual/quien* (línea 25) ya que está en relación con ‘el chico holandés’ (antecedente) y *qué* de pregunta indirecta (línea 28).

Estos errores podrían indicar posibles influencias de la LM suecas en caso de que el aprendiz en sus estrategias asuma que en las líneas 23 y 25 *lo que* podría equivaler a *som* y en la 28 a *vad som* si consideramos que el aprendiz puede hacer traducciones literales de su LM a la LE que estudia (cf. Wiktorsson, 2003).

⁹⁴ En el capítulo 6, figura 3 aparecen sólo las líneas 1-15. Por ello mostramos aquí los ejemplos.

Sería bueno por último apuntar que la combinación *lo que* es la más utilizada por los aprendices suecos del corpus SAELE y por los nativos hispanohablantes del subcorpus del CEDEL2, a la vez que es una locución utilizada en la argumentación más típica de la lengua escrita que de la lengua oral (cf. Escandell, 2008; Domínguez, 2007; López & Borreguero, 2010). Lo anterior es un elemento indicador del nivel en que está situado el aprendiz sueco de E/LE, quien se va acercando a la lengua meta, aunque en ese proceso de la IL cometa errores al utilizar esta combinación de palabras.

7.1.2 Uso de la combinación *de que* por los aprendices suecos de E/LE

La combinación *de que* en sueco tiene dos equivalentes: *att* y *vad*. Aparece en el corpus SAELE en concordancias (n=17) ocupando el segundo lugar en frecuencia de las combinaciones con *que*. Si observamos las palabras antecedentes a *de que* (figura 18) encontramos a sustantivos (n= 7) y verbos (n=2), siendo lo más llamativo para los fines del presente trabajo la frecuencia de aparición de dicha combinación precedida de a pesar (+ *de que*), el hecho (+ *de que*), en caso (+ *de que*) y antes (+ *de que*) en las cuales el elemento precedente es un subordinador. A *de que* la suceden verbos (n=6), artículos (n= 5), pronombres (n=3), adverbios (n=2) y un sustantivo. Se muestran a continuación algunos ejemplos utilizados en los textos del Corpus SAELE:

(59) Todo era maravilloso, ni siquiera *el hecho de que* yo casi no había dormido nada en el avión podía destruir lo que sentí yo en ese momento/Allt var underbart, inte ens *det faktum att* jag inte hade sovit alls på planet kunde förstöra vad jag kände vid det ögonblicket

En el ejemplo 59 notamos que el aprendiz utiliza correctamente la combinación *el hecho de que* y no emplea en este caso el modo subjuntivo en la subordinada sino el indicativo en primera persona del pretérito imperfecto de haber + el participio de dormir: *había dormido*. Como ya hemos mencionado en el capítulo 4, esta combinación puede llevar el verbo en modo indicativo o subjuntivo en dependencia de que la oración subordinada actúe como sujeto o no, y también está en dependencia del conocimiento que tengan los hablantes de lo que se esté tratando.

(60) Está presentada una sifra *de que* unas 350.000 mujeres en España ejercen la prostitución, de ellas sólo unos 30% son de España/En siffrå på *att* cirka 350.000 kvinnor i Spanien är prostituerade har lagts fram, varav endast 30% är från Spanien

DISCUSIÓN DE LOS RESULTADOS

En el ejemplo 60 la combinación *de que* se corresponde con *att* en sueco y aunque ha sido utilizada correctamente, los nativos probablemente elegirían otra combinación y orden en la oración y una alternativa pudiese ser: ‘la cifra presentada indica *que* 350.000 mujeres ejercen la prostitución en España, de las cuales un 30% son españolas’. Este constituye un ejemplo de cómo en la interlengua los aprendices de E/LE utilizan en ocasiones determinadas combinaciones y frases que aunque pueden considerarse ‘no nativas’, son gramatical y semánticamente correctas y se explican en ocasiones como resultado del proceso de intercambio entre la LM y la LE (cf. Laufer, 2000; Odling, 2003; Siyanova & Schmitt, 2008).

(61) Viajé con un amigo y, *a pesé de que* no hemos proyectado algunos cosas antes el viaje, todo el viaje fue muy bien/Jag reste med en kompis och *trots att* vi inte hade planerat några saker innan resan, gick hela resan mycket bra

En el ejemplo 61 el aprendiz ha intentado utilizar la combinación *a pesar de que* en forma ‘casi correcta’, aunque tiene un conocimiento anterior de ‘a pesar de que’ no ha logrado retenerla completa y si observamos el verbo que precede a *de que* y la congruencia en toda la frase se denotan errores ya que debía haber escrito ‘a pesar de que no habíamos planificado algunas cosas antes del viaje’.

El ejemplo ilustra además que en el nivel A2, como se describe en el MCERL, se tiene todavía un repertorio limitado de frases cortas memorizadas (cf. MCERL, cap.3), que el aprendiz suele ser más restrictivo para introducir sus argumentos (cf. Granger, 1998) y a pesar de utilizar combinaciones de palabras que reconoce de situaciones anteriores en que las ha escuchado o utilizado no logra ‘reconstruirlas’ en esta situación comunicativa (cf. Schmitt & Carter, 2004).

(62) Cinco horas *antes de que* tuve que ir al aeropuerto me dieron suero y calmante/Fem timmar *innan* jag åkte till flygplatsen gav de mig en spruta och lugnande medel

En el ejemplo 62 el aprendiz escribe: ‘Cinco horas *antes de que* tuve (Pretérito perfecto simple, indicativo) que ir al aeropuerto me dieron suero y calmante’ pensamos que nos está escribiendo a nosotros para contarnos lo que le sucedió antes de irse al aeropuerto. En otro caso diríamos: ‘Cinco horas *antes de que* me fuera (Pretérito imperfecto, subjuntivo) al aeropuerto me dieron suero y calmante/Cinco horas *antes de que* tuviera (pretérito imperfecto, subjuntivo) que irme al aeropuerto ‘me dieron suero y calmante’ lo que indica, como se plantea en estudios anteriores (cf. Laufer, 2000), la tendencia entre los aprendices a evitar construcciones difíciles, en este caso

el subjuntivo después de la combinación *antes de que*, cuya función es la de ser un subordinador de tiempo.

Se puede apreciar en el mismo ejemplo 62: ‘me dieron suero’ que en sueco podría ser: ‘de gav mig’, que el aprendiz ‘ha copiado’ la frase de su LM ya que un nativo o un aprendiz de nivel más avanzado hubiese escrito la frase: me pusieron/pasaron un suero o sea que se utiliza el verbo *pasar* o *poner* que a su vez serían en sueco *passera* o *sätta*, pero cuando se trata de un suero o una inyección se utiliza el verbo sueco *ge*/dar o *få*/recibir. En el caso de *passera*, ésta se encuentra como 5a. acepción en el diccionario mientras que la primera acepción es entre otras *ge* (las otras son *räcka* [pásame la sal/*räck mig saltet*], *övertämla*, *övertäta*). Mientras *passera* tiene un parecido ortográfico y cierto valor semántico en común con el verbo español, éste es mucho más vago y depende del contexto, mientras que la palabra sueca es de uso mucho más restringido. Con respecto a *sätta* el diccionario incluye cuatro primeras acepciones: *ställa*, *sätta*, *lägga*, *placera*.

Consideramos que esto es un ejemplo evidente de la influencia de la LM (sueco), en la lengua meta.

Para concluir, las líneas de concordancias revisadas permiten ver que los aprendices tienen dificultades en algunos usos de la combinación *de que* en contexto con palabras adyacentes tal como se muestra en algunos de los ejemplos anteriores y se ve específicamente en las líneas de concordancia 4, 8, 9, 13, 14, 15, 16 y 17 (8 de un total de 17 líneas de concordancia), a partir de lo cual consideramos que sería recomendable que los aprendices realizaran actividades donde practicasen las combinaciones *en caso de que*, *antes de que*, *a pesar de que*, *el hecho de que* y *el sentido de que* destacando su posición con respecto a las palabras que los rodea específicamente los verbos que las preceden. Como se ha señalado en el cap. 4 esta combinación se utiliza con subordinadores de tiempo y de condición (cf. Butt & Benjamin, 2000), en conjunciones finales y puede ser parte de una oración subordinada cuyo verbo puede estar en modo indicativo o subjuntivo (cf. Martí Sánchez et al., 2008), cuestiones estas que si se tienen en consideración en la enseñanza implicarán que los aprendices tomen conciencia de las mismas y por consiguiente lleguen a utilizar la combinación de forma similar a como lo hace un nativo de la lengua española.

7.1.3 Uso de la combinación *algo que* por los aprendices suecos de E/LE

Las líneas de concordancia con *algo que* (Figura 5) muestran la combinación precedida por adjetivos (n=3), verbos (n=2), conjunción (n=1), preposición (n=1), artículos (n=1) y en posición inicial (n=1), mientras que el corpus de hispanohablantes (subcorpus CEDEL2) es precedida por verbos (n=5) en 5 de seis líneas (Figura 19).

Como referimos en el capítulo 4, la combinación *algo que* podría tener a *något som*, *vilket*, *vad* y *vad som* como equivalentes en sueco (4.1.2.4). Este hecho nos conduce a discutir dos aspectos con relación al uso de la misma por aprendices suecos de E/LE. El primero es que *algo que* puede utilizarse en lugar de *lo que* y en este caso se correspondería con *vilket* y *något som* en sueco, y el segundo es la relación que se establece entre esta combinación con las palabras adyacentes a ella, específicamente las preposiciones que en sueco no ocupan las mismas posiciones que en español.

Si observamos los ejemplos de las líneas de concordancia con *algo que* del corpus SAELE (Figura 5) y las comparamos con las de los nativos hispanohablantes (Figura 19), veremos que estos últimos prefieren utilizarla en la siguiente combinación: verbo *ser* + *algo que* donde el verbo puede estar en subjuntivo (líneas 2) o en indicativo (líneas 2, 4 y 6) y cuya función es aclarar un hecho o la opinión del hablante.

Los aprendices del SAELE por su parte usan esta combinación con el verbo *ser* en una sola ocasión ya que las líneas 7 y 11 son las mismas. Si comparamos la línea 3 de las concordancias del corpus SAELE con la 5 de las concordancias del corpus CEDEL2 vemos la combinación *algo que* en contexto con *pero* y el verbo *haber* también con una función aclaratoria.

En el corpus de control ARGUS (Figura 28) aparecen las concordancias con *något som* que puede tener función equivalente en sueco a *lo que* en español. Los participantes del ARGUS utilizan también *något som* en las líneas 1, 6, 9, 12, 15, 17, 18, 19, 26, 27, 30 y 32 precedido del verbo *ser* / *vara* en presente *är*/es, de lo cual puede inferirse que los nativos hispanohablantes y los nativos suecos utilizan dicha combinación con la misma función en contexto. Los aprendices suecos de E/LE pueden usar *algo que* en lugar de *lo que* como se ejemplifica en la línea 8 (figura 5) ya que sería mejor haber escrito: ‘...está subiendo, *lo que* puede indicar...’ en lugar de ‘...está subiendo, *algo que* puede indicar...’.

Destaca igualmente la posición en que el aprendiz coloca la preposición *con* en el contexto con *lo que*. Si observamos en la línea 2 de la misma figura ‘Algo que Edward no cuenta con’, se debería haber escrito ‘Algo con lo que Edward no cuenta’. En la línea 4 aparece un ejemplo de uso correcto con una preposición que antecede a *algo que*: ‘el episodio continúa con algo que ocurre’. En el primer caso la utilización incorrecta se debe a la influencia de la LM en la lengua meta como ha sido previamente reportado que ocurre con hablantes anglófonos y de otras lenguas germanas (cf. Söhrman, 2007; Pino, 2006, 2010). Aquí el aprendiz asocia la combinación verbo *räkna*/contar + la partícula *med*/contar que en sueco literalmente sería equivalente a: contar con y la coloca de esa forma en la frase en español.

Quizás no debamos pasar por alto la posible influencia de otras lenguas extranjeras previamente estudiadas (Diagrama 7) sobre todo el inglés⁹⁵ que podría ejemplificarse en el uso de asociaciones o palabras de similitud gráfica como en el caso en que el aprendiz escribe *ocurre* (*ocurrir/occur*). Lo anterior evidencia que en el proceso de aprendizaje de la lengua extranjera se producen también interferencias de otras lenguas adquiridas anteriormente como se ha referido, aunque este aspecto no sea objeto de estudio en esta tesis (cf. Bardel, 2000; Ringbom, 2007; Söhrman, 2007).

Por último se puede apuntar que la combinación *algo que* es la menos utilizada tanto por los hispanohablantes (subcorpus del CEDEL2) como por los aprendices suecos (corpus SAELE).

7.1.4 Uso de la combinación *dice que* por los aprendices suecos de E/LE

El verbo *decir* tanto en la lengua española como en la lengua sueca es un verbo de comunicación (cf. DeMello, 1994; Holm, Nylund et al., 1979). La combinación *dice que*, podría equivaler en sueco a *säger att* o *säga åt någon*⁹⁶ y tiene la peculiaridad de aparecer entre nativos suecos aprendices de E/LE en relación con el uso del modo subjuntivo que, como ya hemos referido en el capítulo 4, suele ser problemático para los aprendices. El problema reside en que deben elegir entre infinitivo, indicativo y

⁹⁵La lengua inglesa comienza a estudiarse en cuarto grado (educación primaria).

⁹⁶*Säga åt någon* no tiene sentido en sueco, pero el significado es más preciso que *dice que*, a saber, *decirle algo a alguien*.

DISCUSIÓN DE LOS RESULTADOS

subjuntivo para el verbo que sigue a *dice que*; y pensar que el subjuntivo se utiliza cuando se expresa cualquier intento, intención o deseo de influir en la acción del sujeto del verbo dependiente. En los ejemplos siguientes: ‘él *dice que* estudias’ (verbo en presente indicativo), ‘él *dice que* estudies’ (verbo en presente subjuntivo) se aprecia claramente las diferencias en formas de usos con independencia de la aparente similitud gráfica.

En las líneas de concordancias con *dice que* del corpus SAELE no se destaca el uso del modo subjuntivo en relación con esta combinación, pero pueden citarse 2 ejemplos en que los aprendices utilizan el imperfecto de indicativo después de *dice que* (Figura 6). En la línea 13 ‘*Dice que* el destino más preferido en España era las Canarias (31,4%), y la segunda era Illes Balears (24,2%)’ y en la línea 14 ‘*Dice que* España era el centro del interés de los turismos’. En estos ejemplos los aprendices deciden utilizar el indicativo (era) en el sentido de *decir que y no decir algo a alguien*, además de que *decir* como verbo de comunicación puede expresar un hecho y en este caso se utiliza el indicativo (cf. Stanley & González, 2007).

Al ver los resultados del uso de la combinación *dice que* llama la atención que ésta no haya sido utilizada en ninguna ocasión por los nativos hispanohablantes y solamente seis por los nativos suecos del corpus ARGUS, lo que vemos como un probable indicador de una baja frecuencia de uso de ésta en la lengua escrita de los nativos de los grupos de control y como muestra de que como aparece en otros estudios los aprendices de LE suelen utilizar combinaciones y palabras del registro oral en sus producciones escritas (cf. Wiktorsson, 2003). El utilizar esta combinación *dice que* podría ser además una estrategia para introducir argumentos con la cual el aprendiz está más familiarizado (cf. Granger, 1998) como por ejemplo en la línea 9: [Dice que Alemania es el país europeo dónde se dan más dinero a los padres mensuales] o en la 2: [Dice que las empresas tienen que tomar la responsabilidad del alojamiento] (corpus SAELE).

A modo de conclusión de éste acápite podemos decir que consideramos que los principales usos y funciones de *que* y las combinaciones con *que* (*lo que*, *de que*, *algo que* y *dice que*) que se aprecian en los textos de los aprendices suecos de E/LE revelan que estos las utilizan en ocasiones según el uso y función asignadas a las mismas por las gramáticas y diccionarios consultados, pero también que con independencia a los errores que cometen son capaces de ‘recrear’ y ‘crear’ combinaciones de acuerdo a su nivel de conocimiento utilizando estrategias de acomodación a partir de su lengua

materna u otras lenguas estudiadas y por ello al comparar los resultados que hemos revisado hasta aquí podemos afirmar que los aprendices producen combinaciones de palabras de acuerdo al principio de elección libre y al fraseológico *que-* al decir de Sinclair- es lo que motiva un *continuum* lingüístico enriquecido.

7.1.5 Correspondencia entre el uso de las combinaciones de palabras con *que: lo que, de que, algo que* y *dice que* entre aprendices suecos de E/LE, nativos hispanohablantes y nativos suecos

Tal y como señalamos en el acápite anterior los aprendices suecos utilizan en ocasiones las combinaciones de palabras con *que: lo que, de que, algo que* y *dice que* de forma equivalente a los nativos hispanohablantes, y en ocasiones en combinaciones fraseológicas que de cierta forma están preelaboradas en su memoria y que ellos rescatan y reconstruyen de acuerdo a la situación comunicativa.

Cabe destacar que se han podido identificar en el presente estudio correspondencias entre el uso de combinaciones de palabras producidas por los aprendices suecos de E/LE (SAELE) y las producidas por nativos suecos e hispanohablantes en los corpus de referencia ARGUS y CEDEL2 que hemos resumido en el cuadro a continuación (Cuadro 21).

DISCUSIÓN DE LOS RESULTADOS

Cuadro 21. Comparación entre corpus

C O R P U S	TOTAL DE PALABRAS	PALABRAS MÁS FRECUENTES	COMBINACIONES MÁS FRECUENTES/ CLUSTERS
S A E L E	37.076	de(1508=4.06%) y (1299=3.50%) la (1211=3,26%) en (1139 = 3,07%) que (1094 = 2.95%)	<i>Que se llama</i> <i>Que me gusta</i> <i>Que es una</i> Creo <i>que es</i> <i>Que es un</i>
C E D E L 2	39.237	De(2.135=5.44 %) <i>Que</i> (1.550=3.45%) La (1.364=3.48%) Y(1144=2,42%) En(1.029=2.62%)	La verdad es Verdad es <i>que</i> En el <i>que</i> Cuenta <i>de que</i> <i>Que</i> he visto
A R G U S	77.650	<i>Att/ que</i> (2.908=3.75%) och/y (2.466=3.18%) <i>det/ el/ la/ lo</i> (2.049=2,64%) <i>i/ en/ dentro de</i> (1.745=2,25%) är/ser (1.713= 2,21%) <i>som/ que/ como</i> (1.652 = 2,13%)	Clusters/ <i>att/ que</i> <i>Att det är/ que es</i> Jag tror <i>att/ creo que</i> Jag tycker <i>att/ pienso, creo, opino</i> <i>que</i> <i>Att det finns/ que existe</i> För <i>att få/ para</i> recibir, obtener Clusters/ <i>som/ que</i> Vad <i>som är/Lo que es</i> Barn <i>som arbetar/</i> Niños que trabajan Är <i>något som/ Es algo que</i> Barn <i>som har/ Niños que tienen</i> De barn <i>som/ Los niños que</i>

En el cuadro 21 se observa que la cantidad total de palabras en los 135 textos escogidos de los respectivos corpus es superior en los corpus de control CEDEL2 (hispanohablantes) y ARGUS (nativos suecos), lo cual probablemente se pueda explicar por una riqueza idiomática superior entre nativos respecto a aprendices y al número de palabras en el texto con independencia que hayamos seleccionado la misma cantidad de textos en los subcorpus de control para que hubiese una homogeneidad respecto al corpus SAELE. Es preciso señalar que para los cálculos de frecuencias no se considera este factor como decisivo gracias a la aplicación de un calculador estadístico que contrarresta este efecto.

La partícula *que* aparece en el corpus SAELE en el quinto lugar (1094 = 2.95%) entre las cinco primeras palabras más frecuentemente empleadas en los tres corpus, mientras que ocupa el segundo lugar en el CEDEL2 (1.550=3.45%). En el ARGUS ocupa su equivalente *att/que* el primer lugar (2.908=3.75%) y *som/que/como* el quinto (1.652 = 2,13%). Los corpus SAELE y el de control CEDEL2 coinciden en tener las mismas cinco palabras de mayor frecuencia aunque éstas ocupen diferentes lugares. Sin embargo hay que señalar que en el corpus de control ARGUS aparece, a diferencia del SAELE y del de control CEDEL2, el verbo *vara/är/ ser/es*, en cuarto lugar.

Al comparar las frecuencias de uso de *que* entre los corpus SAELE con los subcorpus de control CEDEL2 (diagrama 3) y ARGUS⁹⁷ quisiéramos puntualizar que:

Los aprendices suecos de E/LE hacen un infrauso de la partícula *que* respecto a los hispanohablantes del subcorpus del CEDEL2 (Cuadro 16) y la partícula *att* en el corpus de control ARGUS, pero la sobreusan respecto al equivalente sueco *som*.

La partícula *que* aparece en los tres corpus en sus funciones de conjunción (equivalentes en sueco: *att, som*), pronombre relativo (equivalentes en sueco: *som, vilket/vilken/vilka*), comparativo (equivalente en sueco: *mer/mindre...än*), interrogativo (equivalente en sueco: *vad, vilket/vilken/vilka*), exclamativo (equivalente en sueco: *vad, vilket/vilken/vilka*).

La partícula *que* está presente en las cinco combinaciones de palabras/clusters del corpus SAELE y del ARGUS y sólo en cuatro de las del subcorpus de control de

⁹⁷ Las palabras equivalentes a *que*: *vad, vilket, vilken* y *vilka*, cuyas frecuencias de uso son menores, no se han utilizado en esta comparación.

DISCUSIÓN DE LOS RESULTADOS

nativos hispanohablantes CEDEL2. Entre estas combinaciones de palabras/clusters que utilizan los aprendices suecos del corpus SAELE hay dos que aparecen en los corpus de control, *de que* (subcorpus CEDEL2) y tres equivalentes en sueco de *algo que/något som/vad som/vad* (subcorpus ARGUS). *De que* es a su vez, la segunda combinación de palabras más utilizada por los aprendices suecos del corpus SAELE.

Estas co-ocurrencias de tres palabras son interesantes teniendo en cuenta que pueden ser también frases ‘prelaboradas’ por los aprendices y sería importante investigarlas en un estudio posterior para ver la relación léxico-estructura gramatical que se establece en las mismas.

Los aprendices suecos de E/LE del corpus SAELE coinciden en el uso de *y creo que*, *que no es*, *la gente que* y *a pesar de* con los nativos hispanohablantes (subcorpus del CEDEL2) y en *jag tror att/creo que*, *jag tycker att/ pienso que*, *är viktigt att/es importante que* con los nativos suecos (subcorpus ARGUS). La coincidencia de uso en *creo que* y *pienso que* en los tres corpus refleja la similitud entre los tipos de textos elaborados por los participantes quienes expresan sus opiniones utilizando un verbo de creencia y opinión junto a *que*, independientemente que se observan diferencias respecto a las argumentaciones escritas entre nativos suecos e hispanohablantes y aprendices suecos de E/LE debido al nivel idiomático de la lengua española entre estos.

Aunque no colocamos todas las combinaciones de palabras/clusters en este cuadro, debemos decir que obtuvimos un total de 34 en el corpus SAELE, 59 en el subcorpus de control CEDEL2 y 305 combinaciones con *att* y 74 con *som* en el corpus de control ARGUS.

El cuadro 22 (a continuación) compara las frecuencias de usos de las combinaciones en el Corpus de aprendices (SAELE) y el de control de nativos hispanohablantes (CEDEL2), mostrándose que los últimos emplean la combinación *lo que* en frecuencia equivalente al doble de los aprendices y la combinación *de que* casi 4 veces más frecuentemente.

Los aprendices del SAELE emplean la combinación a la inversa ya que *algo que* resulta ser el doble de frecuente en el SAELE respecto al uso de los nativos hispanohablantes (subcorpus CEDEL2), mientras que *dice que* apareció 15 veces en los 135 textos integrantes del SAELE y en ninguno de los del subcorpus de hispanohablantes del CEDEL2, lo que puede indicar que los aprendices en su IL

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

utilizan la combinación *lo que* combinaciones más propias de la lengua oral cuando escriben (cf. Wiktorsson, 2003).

Cuadro 22. Comparación de concordancias de las combinaciones de palabras con *que*. Corpus SAELE y CEDEL2

CORPUS	LO <i>QUE</i>	DE <i>QUE</i>	ALGO <i>QUE</i>	DICE <i>QUE</i>
SAELE	40	17	13	15
CEDEL2	82	62	6	0

Las diferentes combinaciones de palabras en lengua sueca que equivalen en uno u otro contexto fraseológico a *lo que*, *de que*, *algo que* y *dice que* aparecen en el diagrama 4. Esas pueden organizarse en sentido descendiente en cuanto a su frecuencia de aparición: *att/que* (2908); *som/que/como* (1652); *vad/que/lo que/algo que/de que* (275); *vilket/que/lo que/algo que* (50); *det som/lo que* (48); *vad som/lo que/algo que* (46); *något som/algo que/algo que* (33); *vilken/que/lo que* (29); *vilka/que/los que* (18); *den som/algo que* (15); *någon som/lo que/algo que* (14) y *säger att/dice que* (6).

Debido a que los equivalentes suecos de las combinaciones de palabras en español pueden a la vez tener otro significado y función en contexto comparamos solamente aquellos equivalentes de *lo que* en sueco que juegan el mismo papel en ambos idiomas, a saber: *vad som* (46)⁹⁸ y *det som* (48) y el equivalente de *decir que/säga att*, y porque, como muestran los resultados estas combinaciones fueron las más utilizadas por los aprendices suecos respecto a los hispanohablantes del corpus de control CEDEL2 cuando efectuamos la comparación entre las frecuencias de ambos corpus.

La comparación entre las frecuencias de concordancia del SAELE y del ARGUS muestran que los aprendices suecos de E/LE hacen un sobreuso de las mismas como aparece en los cuadros más abajo donde 01= SAELE con la frecuencia de concordancias de 40 (0.11%) y 02=ARGUS con unas frecuencias de concordancias de *det som* de 48 (0.06%) respectivamente *vad som* = 46 (0.06%) que resulta en un sobreuso de *lo que* de 6.57 respecto a *det som*, y de un 7.47 respecto a *vad som*.

⁹⁸ *Vad som /lo que* puede a veces ser sustituido por *algo que/något som* cuando uno traduce del sueco al español según los resultados que se observan en nuestro corpus de control ARGUS.

DISCUSIÓN DE LOS RESULTADOS

tem	O1	%1	O2	%2	LL
Word	40	0.11	48	0.06 +	6.57

Item	O1	%1	O2	%2	LL
Word	40	0.11	46	0.06 +	7.47

- En *dice que/säga att* se muestra también un sobreuso por parte de los aprendices suecos de 13.44%.

O1	%1	O2	%2	LL
Word	15	0.04	6	0.01 + 13.44

En conclusión podemos decir que los aprendices suecos de E/LE (corpus SAELE) utilizan las combinaciones de palabras *algo que* y *dice que* en una proporción mayor que los hispanohablantes del corpus de control CEDEL2, y *lo que/det som/vad som* y *dice que/säger att* con respecto a los nativos suecos del corpus de control ARGUS. Respecto a lo anterior podemos subrayar dos cuestiones, lo primero es que el sobreuso que hacen los aprendices suecos de E/LE de *algo que* respecto a los nativos hispanohablantes y de *lo que* respecto a los nativos suecos podría deberse a que en el discurso escrito en español es común usar *lo cual* cuando la cláusula no es restrictiva (cf. Butt & Benjamin, 2000) y a que *algo que* se puede además utilizar en lugar de *lo que* cuando hace alusión a toda la oración y se corresponde con *vilket* y *något som* en sueco (Véase Cap. 4, ej. 43; cf. Fält, 2000); lo segundo es que el sobreuso de *dice que*, en que se combina un verbo de comunicación + *que* expresa como referimos antes un elemento más típico de la lengua oral que de la escrita, lo cual también se explica a partir de la forma de escribir entre aprendices y nativos, ya que aunque los textos sean argumentativos, los aprendices de SAELE son noveles al encontrarse en un nivel A2-B1 y esto repercute en que su lenguaje sea más restrictivo y repetitivo (cf. Granger, 1998; Wiktorsson, 2003; Scardamalia & Bereiter, 1992).

7.1.6 La influencia de la lengua materna

Para poder indagar acerca de la influencia interlingüística en el proceso de enseñanza aprendizaje de los aprendices suecos de E/LE del corpus SAELE, hemos aplicado en nuestro estudio tanto el ACI, como la LC donde en el ACI – siguiendo a Granger - hemos yuxtapuesto un fenómeno lingüístico (el uso de las combinaciones de palabras con *que*) comparando las construcciones hechas por los aprendices suecos de E/LE (corpus SAELE) con las equivalentes producidas por nativos hispanohablantes (subcorpus de control CEDEL2) y nativos suecos (corpus de control ARGUS), considerando que en este caso el sueco la LM de los aprendices y el español la lengua meta⁹⁹.

Se han tomado en consideración cuestiones tales como el contexto en que se encuentra el aprendiz, la tendencia a evitar producciones difíciles en la lengua meta, la excesiva o poca producción de determinadas estructuras lingüísticas, el principio de ‘acomodar’ frases, y la influencia e interacción con otras lenguas en la adquisición de una nueva y el uso de expresiones más típicas de la lengua oral que de la escrita, analizadas en estudios anteriormente publicados (Cap.3, acápite 3.2).

A partir de lo discutido en los acápites anteriores hemos podido apreciar que los aprendices suecos de E/LE hacen uso de diversos recursos de su LM cuando escriben sus argumentaciones en español. En primer lugar esto se debe a que ellos se encuentran en camino hacia una nueva LE y este camino llamado interlengua está en un proceso de *continuum* interlingüístico donde concurren en nuestro caso la LM de los sujetos que es el sueco y la lengua meta que es el E/LE.

Visto de esta manera se aprecia además mediante los ejemplos explicados anteriormente que la interlengua debe ser entendida como un sistema dinámico en el proceso hacia la nueva LE que no obligatoriamente debe verse afectada de modo negativo por la LM (cf. Pastor Cesteros, 2005), pero en la que fluye continuamente estructuras lingüísticas de la nueva LE, la LM y otras lenguas aprendidas con anterioridad (cf. Baralo, 2004; Cook, 2002; Ringbom, 2007; Sörhman, 2007).

⁹⁹ Véase la propuesta metodológica del ACI de nuestro estudio en el Cap.3, figura 2.

DISCUSIÓN DE LOS RESULTADOS

Los ejemplos denotan la influencia de la LM, y estrategias de las cuales se auxilia el aprendiz de LE para tratar de suplir aquellos conocimientos y habilidades en la lengua meta aún no adquiridos o desarrollados como: el evitar construcciones difíciles en la lengua meta que apenas se emplean en su LM; la acomodación de modos verbales para sustituir la utilización del modo subjuntivo y ajustes en la producción de construcciones de frases y combinaciones de palabras perfectamente comprensibles para el lector, pero que denotan un punto distintivo de lo ‘no nativo’.

Igualmente se observó el sobreuso de combinaciones de palabras tales como *algo que* y *dice que* respecto a nativos hispanohablantes y de *lo que/det som/vad som* y *dice que/säga att* con respecto a los nativos suecos, donde se aprecia el contraste en uso y frecuencia de *dice que* entre aprendices suecos de E/LE y nativos hispanohablantes y suecos.

El siguiente diagrama (5) podría esclarecer mejor esta observación a partir de la interacción entre la LM del aprendiz (sueco) y la LE que estudia (español) donde la IL se ve influenciada de ambas lenguas. Por una parte, en la IL de los aprendices se podría ver el progreso hacia el E/LE en la medida en que estos sobreusan la combinación *lo que* respecto a los equivalentes en sueco nativos suecos del corpus de control ARGUS, si tenemos en cuenta que esta combinación es la más utilizada por los nativos hispanohablantes del subcorpus de control CEDEL2; por otra parte también se evidencia un sobreuso de la combinación *algo que* respecto a los nativos hispanohablantes lo que podría indicar el efecto de la LM (sueco) en la IL ya que esta combinación fue la más utilizada por los nativos suecos del corpus de control ARGUS. Por último, el sobreuso de la combinación *dice que* por los aprendices suecos de E/LE indica, como ya habíamos apuntado anteriormente, la utilización de un combinación de palabras más frecuente de la producción oral en sus producciones escritas lo que parece ser típico de la IL según señalan otros estudios acerca del tema (cf. Granger, 1998; Wiktorsson, 2003).

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

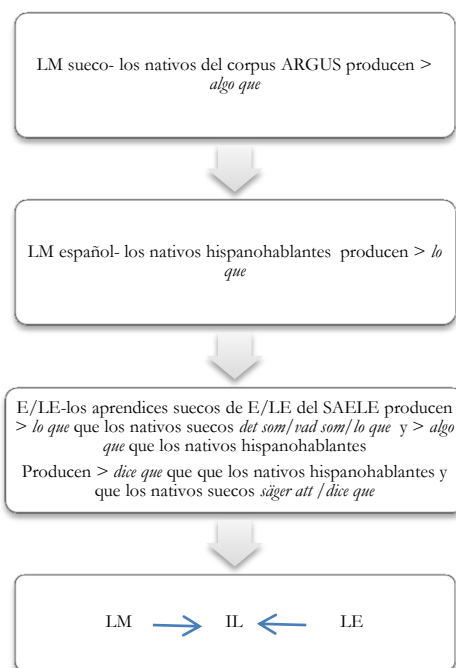


Diagrama 5. Interacción Lengua materna- Interlengua- Lengua extranjera

Una cuestión también interesante que aparece en los ejemplos es la colocación de preposiciones adyacentes a combinaciones de palabras con *que*, en el lugar que las mismas ocupan en la oración en sueco (interacción sueco-español) que parece ser típica también de aprendices anglófonos y de otras lenguas germanas (cf. Söhrman).

Todos estos ejemplos pueden considerarse bien influencias de la LM bien de una LE, bien como parte de estrategias o recursos defensivos de acomodación en el proceso creativo de la interlengua del aprendiz sueco de E/LE participante en este estudio.

7.2 Caracterización sociolingüística del aprendiz sueco de E/LE participante del corpus SAELE

En el capítulo 5 expusimos cómo elaboramos y llevamos a cabo la encuesta para obtener los datos sociolingüísticos de los aprendices suecos de E/LE de este estudio, pero quisiéramos puntualizar aquellos aspectos más relevantes que los sitúan en un contexto social y que pueden influir en su nivel de destreza escrita.

DISCUSIÓN DE LOS RESULTADOS

Para una mejor visualización de la distribución edad/sexo presentamos el diagrama 6, en que se aprecia que la edad promedio de los participantes de la universidad de Gotemburgo (15 mujeres y 14 hombres) era de 28 años mientras que la de los del instituto superior de Jönköping (12 mujeres y 4 hombres) era de 41.

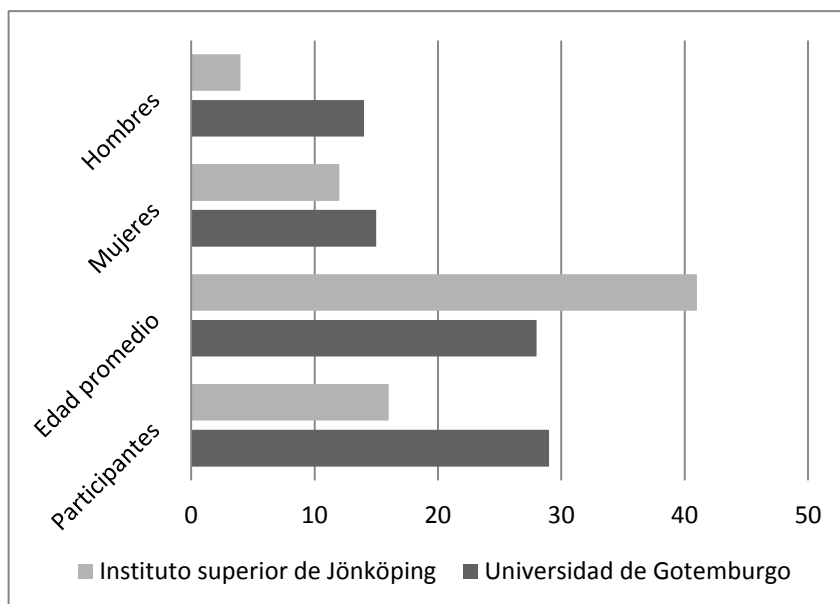


Diagrama 6. Participantes en la encuesta

Las diferencias de edad/sexo mostradas entre los participantes de este estudio con relación al centro de educación superior donde estudian, indican probablemente diferencias demográficas en la población asentada en ciudades grandes y medianas en Suecia, lo cual coincide con la información del ministerio de educación superior sueco en el informe que citamos en el cap.5. Se observa una tendencia de personas más jóvenes y un casi igual número de sujetos en relación hombres y mujeres en la universidad de Gotemburgo que en el instituto superior de Jönköping. En este último la muestra era de 4 hombres y 12 mujeres mientras que en la universidad de Gotemburgo era de 14 hombres y 15 mujeres.

Los cuadros 7 y 8 (Cap.5) indican que el tiempo previo de estudio de E/LE entre los sujetos procedentes de ambos centros de educación superior fue aproximadamente de 3 años, el 70% de los sujetos de la universidad de Gotemburgo en el curso lectivo 2008/09 respectivamente el 50% del instituto superior de Jönköping. Por otra parte, el 25% de los matriculados en el curso 2008/09 respectivamente el 50% de los

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

correspondientes al curso 2009/10 en el mismo orden cursaron estudios en un país hispanohablante. El 64,4 % refirió además haber cursado estudios en Suecia.

En el diagrama 7, aparece recogida gráficamente la correlación entre el tiempo de estudio de E/LE, el tiempo de estudio en un país hispanohablante y en Suecia por universidades.

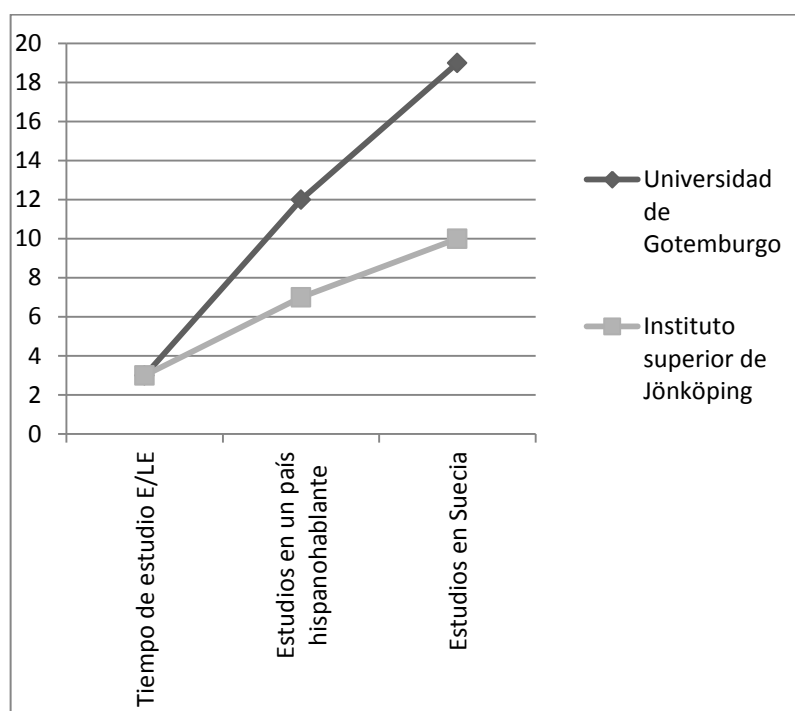


Diagrama 7. Relación entre tiempo de estudio de E/LE en un país hispanohablante o en Suecia

Aquí podemos generalizar que la mayoría de los participantes había cursado estudios de E/LE en Suecia al menos durante tres años y que aproximadamente la mitad de ellos había realizado una estancia en un país hispanohablante, lo que es un indicador que podemos asociar con el nivel de destreza escrita que se observa en los textos de los participantes teniendo en cuenta que cuando se estudia una LE en un país donde es la lengua materna esto posibilita que el aprendiz tenga un contacto directo con los nativos.

DISCUSIÓN DE LOS RESULTADOS

Otro parámetro sociolingüístico de la encuesta se refiere al conocimiento de lenguas extranjeras que poseen los participantes. En el diagrama 8 se presenta la relación de lenguas extranjeras que los participantes refirieron haber estudiado.

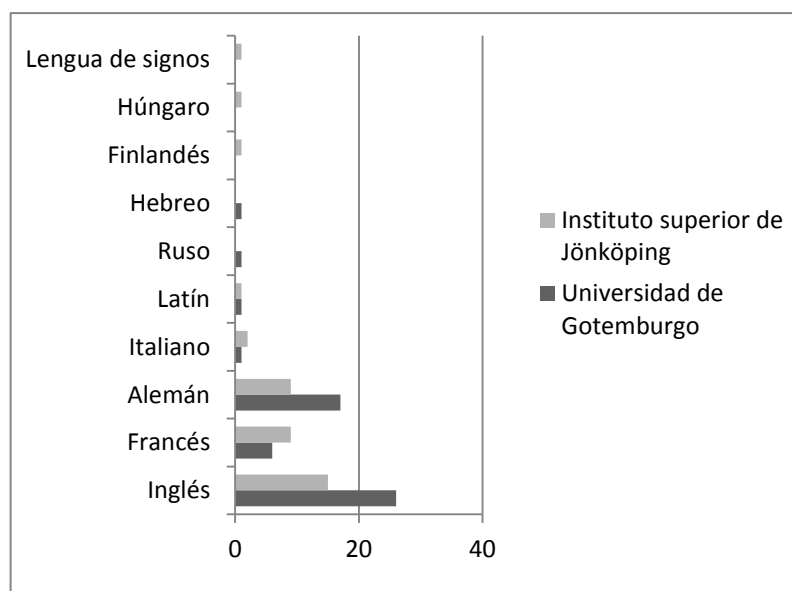


Diagrama 8. Otras LE estudiadas por los participantes

Los de la universidad de Gotemburgo habían estudiado: inglés, francés, alemán, italiano, latín, ruso y hebreo, mientras que los de Jönköping habían estudiado: inglés, francés, alemán, italiano, latín, finlandés, húngaro y la lengua de signos. La cantidad de aprendices entre ambos centros fue diferente, pero se evidencia la preponderancia del inglés, alemán y francés, destacándose además que la mayoría de los participantes de la universidad de Gotemburgo había estudiado alemán, mientras que los del instituto superior de Jönköping muestran una equivalencia relativa entre el alemán y el francés. Para nuestro estudio este resultado es interesante porque nos muestra que los aprendices suecos de E/LE supuestamente poseen estrategias de aprendizaje de LE lo que contribuirá a que sea más fácil el aprender una nueva lengua.

La autovaloración de los aprendices suecos de E/LE de sus destrezas en la comprensión de lectura y expresión escrita aparece resumida en el diagrama 9.

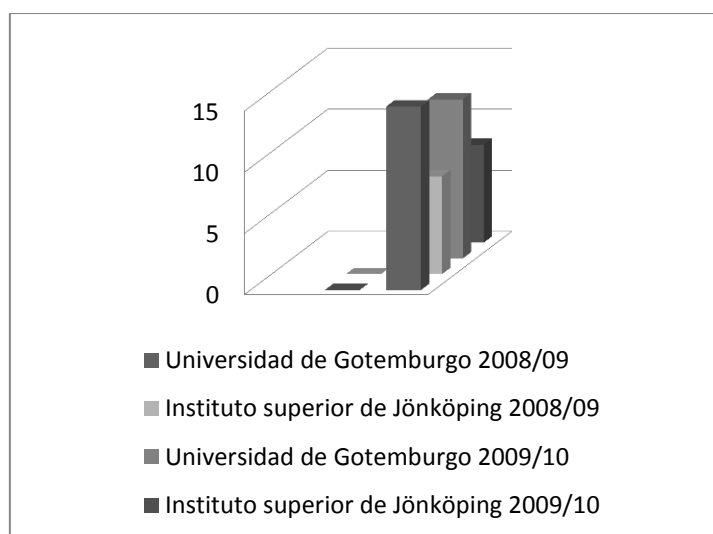


Diagrama 9. Autovaloración de las destrezas escritas

Se observa que los aprendices suecos de E/LE de ambos cursos lectivos (2008/09; 2009/10) de la universidad de Gotemburgo valoran sus destrezas a un nivel más alto que los aprendices suecos de E/LE que estudiaban en el instituto superior de Jönköping. Quizás el resultado anterior se deba al hecho de que los aprendices suecos de E/LE de la universidad de Gotemburgo sean más jóvenes y seguros de sí mismos y que los de más edad por experiencia se manifiesten con más cautela al emitir criterios valorativos; a eventuales diferencias en cuanto a los tipos de cursos en que participaron si tenemos en cuenta que el curso de la universidad de Gotemburgo era a tiempo completo y más intensivo que el de Jönköping (véanse los Anexos) u a otros factores posibles que no son objeto de estudio en este trabajo.

En síntesis, la encuesta realizada nos han permitido conocer mejor qué factores contextuales podrían influir en el nivel de su destreza escrita, además que vemos a los aprendices ‘situados’ según la teoría sociocultural y no sólo como individuos ‘productores de textos’ (cf. Säljö, 2000).

Los resultados obtenidos nos ofrecen elementos de apoyo para presentar un perfil del aprendiz sueco del corpus SAELE e incluso esclarecer cuál es la correspondencia entre los descriptores del MCERL para el nivel A2-B1 y la competencia real mostrada por los aprendices suecos de E/LE participantes en este estudio ya que como planteamos anteriormente hemos pretendido abordar la relación entre ‘saber hacer’ de la destreza escrita y las competencias lingüísticas que influyen en el ‘saber

crear²- según lo planteado por el MCERL y la propuesta de González Piñeiro et al. (2010). (Cap.3).

7.3 Correspondencia entre los descriptores del MCERL y la competencia de los aprendices suecos de E/LE participantes en el estudio

El corpus SAELE de aprendices suecos de E/LE se recopiló en la universidad de Gotemburgo y el instituto superior de Jönköping, 2 de los 16 centros de educación superior de Suecia donde el presente estudio se efectuó entre 2008/2010 (Acápite 3.3, capítulo 3). Los estudiantes que acceden a estos cursos deben poseer un nivel A2 de acuerdo al MCERL (Instituto Cervantes), en el cual el aprendiz debe tener un repertorio de elementos lingüísticos básicos que le permita abordar situaciones cotidianas de contenido predecible, aunque adapte el mensaje y busque palabras para producir expresiones breves y habituales con el fin de satisfacer necesidades de comunicación sencillas y concretas: datos personales, acciones habituales, carencias y necesidades, demandas de información. Pueden en este nivel utilizarse estructuras sintácticas básicas y comunicarse mediante frases, grupos de palabras y fórmulas memorizadas al referirse a sí mismos y a otras personas, a lo que hace, a los lugares, a las posesiones. Además tienen un repertorio limitado de frases cortas memorizadas que incluyen situaciones predecibles de supervivencia e incurrir en malentendidos e interrupciones si se trata de una situación de comunicación poco frecuente.

En el nivel B1 se supone que el aprendiz posea suficientes elementos lingüísticos para describir situaciones impredecibles, explicar los puntos principales de una idea o un problema con razonable precisión y para expresar pensamientos sobre temas abstractos o culturales tales como la música y las películas, al mismo tiempo que posea suficiente vocabulario para expresarse con algunas dudas y circunloquios sobre temas como la familia, aficiones e intereses, trabajo, viajes y hechos de actualidad, aunque las limitaciones léxicas provoquen repeticiones e incluso, a veces, dificultades en la formulación.

Los textos recopilados en SAELE poseen una dimensión comunicativa, pragmática o estructural¹⁰⁰ dependiendo del nivel en que se encuentra el aprendiz sueco de E/LE que lo haya escrito (entre A2 y B1) y se corresponden con las competencias lingüísticas, sociolingüísticas y pragmáticas referidas en el MCERL (véase cap.3, acápite 3.4.1) y por sus temas, el sentido de lo que quieren comunicar, la estructura que eligen para hacerlo, las herramientas textuales empleadas, y el resultado muestra precisamente una competencia léxica, sintáctica, gramatical y pragmática que muestran una tendencia general de desarrollo del nivel A2 hacia el B1 de acuerdo con nuestro estudio.

En los cuadros 1 y 2 aparece una adaptación que hemos realizado a la propuesta metodológica de González Piñeiro et al. (2010) (véase el Cap.3) para explicar más detalladamente la correlación entre las destrezas/competencias en la lengua española. El análisis de los 135 textos escritos examinados según esta metodología permite plantear que una parte de los aprendices suecos de E/LE (corpus SAELE) han sido capaces de reunir información y estructurarla en forma aceptable antes de enviarla al profesor como puede constatarse en los ejemplos de textos presentados en el capítulo 6 (Cuadros 11-14), ello demuestra que el aprendiz se vale de estrategias del tipo *know how* que les permite ascender sucesivamente en la escala de dominio de las técnicas de la composición (cf. Martínez Sánchez, 2004b; Maqueo, 2005), pero también indican la necesidad de reforzar la enseñanza no solamente en cuanto a las técnicas de redacción y la reflexión del producto final ya que como muestran algunos investigadores el proceso de construir un texto precisa del desarrollo de habilidades en la preparación del texto, o sea que se requiere de un tipo de ‘enseñanza pre-textual’ (cf. Scardamalia & Bereiter, 1992).

Los integrantes de las muestras de nuestro estudio participaban en los cursos de destrezas que integraban respectivamente el curso *Spanska 1-30 hp* (Jönköping) y el *Grundkurs I 1-30 hp* (Gotemburgo), los cuales, aunque como asignatura no ocupaban el mismo lugar en ambos centros de educación superior, tuvieron por sus contenidos un enfoque similar (comunicativo centrado en el alumno) y un método enfocado en la combinación del trabajo individual con el de parejas y grupos, estricta planificación de temas y el desarrollo de tareas de comprensión oral y escrita orientadas al desarrollo de las destrezas de los aprendices en E/LE.

¹⁰⁰ Véase sobre las dimensiones textuales en Loureda Lamas (2009: 20).

DISCUSIÓN DE LOS RESULTADOS

En este estudio se pone de relieve que si bien el aprendiz de E/LE no dispone en la producción del lenguaje escrito de elementos del lenguaje extraverbal, se auxilia de frases prefabricadas y combinaciones de palabras a veces extraídas de su LM u otras LE, y del uso a veces excesivo de *que* y sus combinaciones como recurso de apoyo para aclarar y argumentar (cf. Carrió Pastor & Seiz Ortíz, 1999; Wiktorsson, 2003).

Como ejemplo puede emplearse el cuadro 11 donde aparece una redacción argumentativa donde el aprendiz quiere comunicar al lector sus opiniones acerca de un artículo periodístico, disponiendo estructuralmente el texto en una introducción pequeña y se presentan opiniones en oraciones y frases relativamente largas enlazadas por los conectores *pero*, *porque*, *aunque* y *que* (como conjunción, pronombre y comparativo), así como combinaciones tales como *mientras que*, *a pesar de que*, *claro que*, *algo que*, etc. Errores de vocabulario y gramaticales tales como: *el resulto*; *vivir en pobreza* se registran igualmente.

En el propio texto puede leerse: ‘...La figura justifica los datos y la información *que* encontrábamos en el artículo. En la figura, se ve muy claro *que* México tiene un coeficiente de Gini mucho más alto *que* el resto de los miembros de OCDE, lo más alto el coeficiente de gini tanto más desigualdad en el país, es decir que tienen una desigualdad más alta...’ Este fragmento es un ejemplo que tipifica la escritura en la interlengua del aprendiz que se esfuerza por reconstruir una combinación: ‘lo más alto el coeficiente de gini tanto más desigualdad en el país, es decir que tienen una desigualdad más alta’

El aprendiz conocía probablemente la combinación ‘*mientras más –más*’ con anterioridad y aquí se nota una interacción entre el ‘saber hacer’ y el ‘saber crear’ en la interlengua.

En el texto del cuadro 12 se muestra un ejemplo de competencia léxica, sintáctica, pragmática y gramatical correspondiente al nivel A2 del MCERL. La idea del tema elegido se plasma en un texto sencillo, dividido en introducción, desarrollo y desenlace, donde el aprendiz comunica/cuenta su paseo con auxilio de oraciones cortas enlazadas por *y*, *que* y *pero*, y utiliza correctamente algunos tiempos verbales del indicativo. Como refieren algunos autores los aprendices noveles suelen elaborar este tipo de textos con oraciones cortas (cf. Scardamalia & Bereirter, 1992).

El texto del cuadro 14 cumple la dimensión comunicativa donde el lector (yo) puede seguir y entender lo que el aprendiz ha escrito. El texto tiene una buena disposición.

El aprendiz introduce la película, después desarrolla el contenido y al final la recomienda. Aparecen oraciones largas y cortas entrelazadas con *y*, *que*, *porque* y *pero*. Se pueden apreciar errores gramaticales, de sintaxis y del léxico como: ‘Tiene una enfermedad *que* la va a hacer ciega; A pesar de *que* ella tiene un tiempo duro, siempre está en buen humor y es muy concederado ; La película es ruedo con un cámara temblado para reflecta la realidad mejor’ que no solamente permiten clasificar el nivel de destreza/competencia del aprendiz (en redacción y riqueza de palabras), sino también identificar elementos de la IL donde lo distintivo de lo ‘no nativo’ se evidencia en la construcción de algunas expresiones.

Estos hechos ratifican que el aprendiz se encuentra también en un nivel A2. La diferencia entre las dos producciones argumentativas plasmadas no radica solamente en la cantidad de errores cometidos ni conectores utilizados sino en la simultaneidad de la interacción entre las dimensiones comunicativas, estructurales y pragmáticas y la competencia global del producto final elaborado: el texto.

Estos tres ejemplos de textos argumentativos extraídos del corpus SAELE ilustran la tendencia en el nivel A2 hacia el B1 (MCERL) de los integrantes de las muestras del estudio a través de la variación de temas escogidos, la capacidad de argumentación, la utilización de oraciones cortas y largas; frases sencillas e idiomáticas, el dominio de estructuras basadas en introducción, desarrollo y conclusión, el empleo de elementos de enlace y conectores como: *y*, *que*, *pero*, *lo que*, *a pesar de que*, *claro que*.

Expresado en otros términos, se demuestra en el corpus SAELE la presencia de errores como parte inevitable del proceso de aprendizaje. Los aprendices suecos de E/LE muestran ‘saber crear’ a partir de lo que ‘saben hacer’, pero no hay que dejar de considerar la necesidad de tratar de estimular el proceso de trabajo ‘antes de entregar el texto’ o estrategias lingüísticas del tipo *know-how* (saber cómo o dominio del método) que abarquen tanto el trabajo en el aula de E/LE como el trabajo individual en casa, porque este hecho ayuda a que el alumno inmerso en la solución del problema concientice sus dificultades.

7.3.1 Perfil del aprendiz sueco de E/LE participante en el estudio

A continuación sintetizamos los principales elementos que caracterizan al aprendiz sueco de E/LE participante en nuestro estudio:

- Competencia situada en nivel A2 hacia el B1 (MCERL), teniendo en cuenta las diferencias individuales observadas en los textos producidos y guiándonos por los descriptores del MCERL y la adaptación de la propuesta de González Piñeiro & et al. (2010).
- Edad promedio de 28 años en la universidad de Gotemburgo y 41 en el instituto superior de Jönköping lo que muestra una correspondencia con el informe emitido por el ministerio de educación superior sueco de 2010. Lo que significa que los aprendices suecos de la universidad de Gotemburgo que formaron parte de este estudio eran más jóvenes que los del instituto superior de Jönköping.
- El 64 % de ellos habían cursado estudios en Suecia alrededor de tres años y un 25% del curso 2008/09 respectivamente el 50% de los correspondientes al curso 2009/10 en el mismo orden cursaron estudios en un país hispanohablante. Esto nos indica que la mayoría de los participantes había realizado sus estudios aproximadamente tres años en Suecia y que casi la mitad había estado en una estancia en un país hispanohablante, lo cual se refleja en las diferencias en la competencia que muestran los textos del corpus SAELE.
- La mayoría había estudiado otras lenguas extranjeras donde se destaca en primer lugar el inglés, seguido del alemán y el francés. Este resultado nos muestra que los aprendices suecos de E/LE poseen experiencias en el estudio de una LE y esto repercute en el proceso de aprendizaje del E/LE.
- Producen un lenguaje escrito en ocasiones influidos por la LM (sueco), que les posibilita en ocasiones estrategias para compensar déficits de conocimientos/habilidades en la lengua meta.
- ‘Acomodan’ y ‘reconstruyen combinaciones de palabras’, tratando de evitar construcciones de la lengua meta poco habituales en la LM y por tanto difíciles de apropiarse en el proceso de aprendizaje de la lengua meta.

- Emplean combinaciones de palabras según el principio idiomático aunque producen en ocasiones expresiones que si bien resultan comprensibles en la lengua meta, difieren sustancialmente de las elegidas por un nativo hispanohablante lo que le confieren un toque distintivo de ‘lo no nativo’.
- Utilizan combinaciones de palabras/clusters que coinciden con las utilizadas por los nativos hispanohablantes: *y creo que*, *que no es*, *la gente que* y *a pesar de*. Asimismo otras que coinciden con las de los nativos suecos del corpus de control: *jag tror att / creo que*, *jag tycker att/ya opino que*, *är viktigt att/es importante que*. Las combinaciones de palabras que coinciden en los tres corpus están relacionadas con el tipo de texto que escriben los sujetos. Por ejemplo: ‘*y creo que*’ (corpus SAELE) respecto a ‘*y creo que*’ (subcorpus CEDEL2) y ‘*Jag tror att/creo que*’ (subcorpus ARGUS). En estas combinaciones hay un verbo de opinión (creer) precediendo a *que* en la argumentación.
- Hacen un infrauso de la partícula *que* respecto a los hispanohablantes del subcorpus del CEDEL2 y la partícula *att* en el corpus de control ARGUS, pero la sobreusan respecto al equivalente sueco *som*.
- Sobreuso de las combinaciones de palabras: *lo que/det som/vad som* y *dice que/säger att* con respecto a los nativos suecos (corpus de control ARGUS) y de *algo que* y *dice que* respecto a los nativos hispanohablantes.
- El sobreuso de *dice que* indica la utilización de combinaciones más propias de la lengua oral en la lengua escrita.
- Escritura creativa de la interlengua. Esto se fundamenta en que de acuerdo a los textos que forman parte del corpus SAELE (algunos de los cuales analizamos en este capítulo) los aprendices suecos de E/LE han demostrado saber construir un texto utilizando recursos y experiencias tomados del propio proceso de enseñanza-aprendizaje en el aula universitaria, la LM, otras LE e incluso de la estancia en un país hispanohablante.

CAPÍTULO 8. CONCLUSIONES

El objetivo general que nos propusimos en este estudio fue comparar el uso de algunas combinaciones de palabras de la producción escrita de aprendices suecos de E/LE (nivel A1-B1 de acuerdo a la escala del MCERL) con las producidas por nativos hispanohablantes y suecos, utilizando un enfoque de análisis contrastivo de la interlengua basado en la lingüística de corpus. Este objetivo general lo delimitamos en cinco objetivos específicos (véase el Cap. 1) que una vez contestados en la discusión nos permitieron presentar una propuesta de perfil del aprendiz sueco de E/LE participante en este estudio.

Los hallazgos aportados permiten sacar conclusiones en dos sentidos. En primer lugar con referencia a la funcionalidad de *que* y las combinaciones con *que* estudiadas y en segundo lugar con referencia al uso que hacen los aprendices del mismo. Veamos:

- Hemos podido constatar que *que* es utilizado de acuerdo a su funciones idiomáticas en el corpus SAELE , en el CEDEL2 (subcorpus de nativos hispanohablantes), así como sus equivalentes en lengua sueca *att/som/vilken/vilket/vilka/vad/än* en el ARGUS (subcorpus de nativos suecos).
- En cuanto a la frecuencia de uso, *que* (n=1.094, 2.95%) ocupó el quinto lugar en el Corpus SAELE y el segundo lugar en el subcorpus de nativos hispanohablantes CEDEL2 (n=1.550, 3.45%). En el subcorpus de control de nativos suecos ARGUS el equivalente sueco *att* (n=2.908, 3.75%) ocupó el primer lugar seguido del equivalente *som* (n=1.652, 2.13 %). Pudimos por tanto deducir que se emplea con menos frecuencia en el SAELE que en el CEDEL2 y en el ARGUS respecto a la partícula *att*, pero hay un sobreuso relativo con relación a *som*.

Pasamos ahora a ver el uso que se hace de las distintas combinaciones de *que*.

- Los nativos hispanohablantes utilizan más frecuentemente la combinación *lo que* (n=82), mientras que los nativos suecos prefieren la combinación *algo que* cuando esta equivale a *något som* (n=33) y *vad som* (n=46).

- Los nativos hispanohablantes utilizan la combinación *lo que* (n=82) con el doble de frecuencia y *de que* (n=62) casi 3 veces más frecuentemente que los aprendices suecos de E/LE. Estos últimos, en cambio, emplearon las combinaciones *algo que* y *dice que* un 0.4 % y un 21,6% más frecuentemente que los hispanohablantes y *lo que/det som/vad som* y *dice que/säger att* más frecuentemente que los nativos suecos. *Lo que* tuvo un 6.57% de sobreuso respecto a *det som* y de 7.47% respecto a *vad som* y *dice que* presentó un sobreuso de 13.44 % respecto a su equivalente *säger att*. El sobreuso de *algo que* respecto a los nativos hispanohablantes y *lo que* respecto a los nativos suecos muestra el *continuum* en la IL de los aprendices suecos de E/LE si tenemos en cuenta que la combinación *algo que* es la más utilizada por los nativos suecos (subcorpus ARGUS) y *lo que* la más utilizada por los nativos hispanohablantes (subcorpus CEDEL2). El sobreuso de la combinación *dice que* respecto a los resultados de los corpus de control (subcorpus ARGUS y subcorpus de nativos hispanohablantes CEDEL2) por parte de los aprendices suecos de E/LE que coinciden con los estudios efectuados anteriormente indica que los aprendices suecos utilizan combinaciones más propias de la lengua oral en su expresión escrita.
- Los aprendices suecos de E/LE utilizan las combinaciones de palabras con *que* en ocasiones de la misma forma que los nativos hispanohablantes, pero también cometen errores en su uso o en el contexto en que se valen de *que* con otras palabras adyacentes a estas combinaciones. De ello se deduce que no es suficiente trabajar con combinaciones y agrupaciones de palabras, sino que también hay que ver estas combinaciones como parte de un contexto lingüístico en que el léxico y la gramática están en interacción.
- Hemos asimismo podido ver cómo el aprendiz a veces ‘evita’ estructuras ‘difíciles’, esto es, aquellas que apenas se emplean o no existen en su LM (en nuestro caso, el sueco) y en otros ‘acomoda’ determinadas combinaciones de palabras o modos verbales. Un caso sintomático es el uso del modo subjuntivo en el contexto fraseológico con las combinaciones que estudiamos, donde hemos podido constatar que el aprendiz ‘evita’ usar este modo en algunos casos, o cuando se vale de combinaciones de palabras y frases de modelo estructural parecido al de su LM cuyo resultado lingüístico deviene en expresiones comprensibles para el lector, pero evidentemente matizadas por ‘lo no nativo’ resultan en consecuencia menos idiomáticas a los oídos del hispanohablante.

CONCLUSIONES

- A lo anterior se suma la estrategia de lo que denominamos ‘reconstrucción de combinaciones de palabras’, cuando el aprendiz trata de incorporar combinaciones de palabras que bien ha trabajado en el aula anteriormente o bien ha leído o escuchado en alguna conversación, pero no memoriza completamente, y que decide insertar al texto aunque estas no sean las opciones idiomáticas preferidas por un hablante nativo.
- Los textos elaborados por los aprendices suecos de E/LE indican que su nivel de competencia léxica, sintáctica, gramatical y pragmática se corresponde a un nivel A2 hacia el B1 (MCERL), destacándose que se aprecian diferencias individuales entre los participantes debido a factores como la edad, el tiempo de estudio de E/LE en Suecia o un país hispanohablante, el estudio de otras LE y las habilidades en las tareas de producción escrita.

La base de las preguntas de investigación se sustentaban en la hipótesis de que los aprendices suecos de E/LE utilizan con una mayor frecuencia determinadas combinaciones de palabras con *que*: *lo que*, *de que*, *algo que*, *dice que* con respecto a las correspondientes utilizadas por los nativos suecos e hispanohablantes de los corpus de control, y esto estaría en relación con la influencia de la lengua materna (sueco) y a estrategias de acomodación de las cuales ellos se auxilian en su interlengua. Igualmente que las diferencias en formas y frecuencias de esas combinaciones en comparación con los hispanohablantes van disminuyendo según el aprendiz adquiere mayores conocimientos y desarrolla una mayor habilidad de destrezas en la lengua meta.

Con respecto a la primera parte de la hipótesis es importante apuntar que ésta se cumple parcialmente dado que, como expusimos, hemos podido constatar un sobreuso de algunas combinaciones de palabras con *que* por parte de los aprendices suecos de E/LE respecto a los nativos participantes de los corpus de control, aunque el sobreuso no es de todas las combinaciones de palabras con *que* en relación con los nativos hispanohablantes y suecos. Por otra parte, si comparamos los resultados de este estudio con los obtenidos en uno de los realizados con anterioridad en que participaban aprendices de E/LE de un nivel superior (B1-B2), se constata que los de este estudio (A2-B1) hicieron un sobreuso de *que* y sus combinaciones respecto a los aprendices suecos de nivel avanzado.

Consideramos oportuno justificar también la relevancia de este estudio tanto en lo que concierne a aspectos de la LC, el ACI como a la didáctica en la enseñanza-aprendizaje del E/LE para aprendices suecos situados en un nivel A2-B1:

- El ACI basado en la LC es una buena combinación para indagar las variaciones en el uso de la lengua, especialmente en la de aprendices de LE. Este modelo permite contrastar uno o varios aspectos de la interlengua del aprendiz que sea objeto de estudio con sus equivalentes en la lengua materna y la lengua meta extranjera de forma que el investigador se percate más fácilmente de los errores, las interferencias de la LM y otras lenguas adquiridas anteriormente y a que incluso pueda comparar la interlengua de aprendices que estudien una LE, pero cuya LM sea diferente.
- La colección de textos electrónicos de aprendices suecos de E/LE son una fuente importante para descubrir qué caracteriza su escritura, los recursos que utilizan en la comunicación, los errores y las combinaciones de palabras que usan más frecuentemente. Esto se sustenta en que esta compilación (que además también podría ser de lengua oral) nos ofrece la oportunidad de trabajar con un material auténtico y fiable (nunca debemos perder de vista seguir los principios metodológicos y éticos en la elaboración del corpus) que podemos utilizar en investigaciones posteriores, permite mayor credibilidad en los resultados cuantitativos y por último se debe puntualizar que el procesamiento lingüístico del material con ayuda de la computación nos ahorra tiempo y permite una mejor organización de la muestra de estudio.

Después de referirnos a los principales aportes de esta investigación, quisiéramos comentar que nuestra investigación abre a nuevos campos que sería sumamente interesantes indagar en un futuro próximo. He aquí algunos ejemplos de áreas de investigación a desarrollar según los resultados del presente estudio:

- Sería conveniente realizar un etiquetamiento adecuado del corpus SAELE-para facilitar utilizarlo en comparaciones, y así por ejemplo poder descubrir similitudes y diferencias entre aprendices suecos de E/LE y otros aprendices de español con otra lengua materna.
- Ampliar el corpus SAELE con otras categorías de participantes que estén situados en otros niveles de aprendizaje según el MCERL.

CONCLUSIONES

- Incorporar al corpus SAELE un subcorpus de textos de nativos suecos escritos en LM para tenerlo como corpus de control en futuros estudios (teniendo en cuenta los mismos requisitos y temas de las argumentaciones de los aprendices suecos de E/LE).
- Investigar otras combinaciones de palabras recurrentes en el corpus SAELE ya que este estudio nos ha demostrado que mediante el análisis de combinaciones de palabras podemos llegar a descubrir cuáles son las mayores dificultades para un aprendiz sueco de E/LE.
- Desarrollar material didáctico que permita trabajar con combinaciones y agrupaciones de palabras en contexto, teniendo en cuenta que la fluidez textual depende en buena medida del dominio en el uso y memorización de éstas combinaciones y que es precisamente con el principio idiomático que los aprendices de LE tienen más dificultades.
- Fortalecer el desarrollo paralelo de todas las destrezas de la lengua, pero hacer una distinción entre aquellos elementos conectores más utilizados en el registro oral que en el escrito, dado que hemos podido constatar que los aprendices suelen en ocasiones producir sus textos utilizando combinaciones que son más propias del registro oral.
- Planificar adecuadamente el trabajo de producción escrita desde el momento inicial (fase de ‘antes de comenzar a escribir o pre-textual’) hasta el final (‘la reflexión sobre cómo he escrito’) para, de esta forma, poder llegar a presentar un producto final mejor elaborado desde el principio hasta el final.
- Aunque estamos de acuerdo en que la lengua española es el vehículo principal en la enseñanza, sería igualmente importante esclarecer desde el punto de vista contrastivo los aspectos donde los aprendices suecos suelen cometer errores como consecuencia de la influencia de su lengua materna. Esto es consecuencia de que, según expusimos, el ‘evitar construcciones difíciles’ en la intelengua es evidentemente un recurso muy empleado por los aprendices.

A modo de colofón, vale resumir diciendo que, aunque con este estudio hemos dado respuesta a algunas de las inquietudes que teníamos antes de comenzar, quedan aún muchas interrogantes a las que responder y caminos investigativos a explorar teniendo como herramientas al ACI basado en corpus. Esperamos que la presente investigación haya dado buena muestra de ello y de que, por tanto, sea una brecha de investigación que muchos quieran seguir tomando para de esta forma llegar a tener

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

un conocimiento más amplio del arduo camino que implica el aprendizaje de lenguas
extranjeras.

BIBLIOGRAFÍA

- Aijmer, K. (2002). *English Discourse Particlar. Evidence from a Corpus*. Amsterdam: John Benjamins Publishing Company.
- Aijmer, K. & Stenström A. (2004a). *Discourse Patterns in Spoken and Written Corpora*. Göteborg University / Bergen University: John Benjamins Publishing Company.
- Aijmer, K. & Altenberg, B. (2004b). *Advances in Corpus Linguistics*. Papers from the 23rd International Conference on English Language research on Computerised Corpora (ICAME 23). Amsterdam/Philadelphia: Rodopi
- Aijmer, K. (2006). *Pragmatic Marker in Contrast*. Amsterdam: Elsevier Science & Technology.
- Aijmer, K. (2009). Introduction: Corpora and language teaching. En K. Aijmer, (Ed.), *Corpora and Language Teaching*, 1–10. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Alarcos Llorach, E. (2005). *Gramática de la lengua española*. Madrid: ESPASA.
- Alarcos Llorach, E. (1994). Español “QUE”. En Alarcos Llorach, E., *Gramática Funcional del español*, 260-274. Madrid: Gredos.
- Alcón Soler, E. (2002). *Bases lingüísticas y metodológicas para la enseñanza de la lengua inglesa*. Castelló de la Plana: Publicacions de la Universitat Jaume I.
- Allén, S. (1990). *Norstedts svenska ordbok - 80.000 ord och fraser*. Norbok a.s, Oslo/Gjøvik: Norstedts Förlag AB
- Alonso Ramos, M. (2007). El corpus lingüístico en la enseñanza del español como LE. *Boletín de ASELE* 37, 11-27. Disponible en: <http://formespa.rediris.es/pdfs/asele37.pdf>. 22-12-2010.
- Altenberg, B. (2002a). Causative constructions in English and Swedish. A corpus-based contrastive study. En B. Altenberg & S. Granger (Eds.), *Lexis in contrast. Lexis in Contrast. Corpus-based approaches*, 97-116. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Altenberg, B. (2002b). Concessive connectors in English and Swedish. En H. Hasselgård, S. Johansson, B. Behrens & C. Fabricius-Hansen (Eds.), *Information Structure in a Cross- Linguistic Perspective*. Amsterdam: Rodopi.
- Altenberg, B. (2007). The Correspondence of Resultive Connectors in English and Swedish. *Nordic Journal of English Studies*, Volume 6, No.1, 1-26.
- Altenberg, B. (2010). Conclusive English 'then' and Swedish 'då'. A corpus-based contrastive study. *Languages in Contrast* 10, 102-123.
Amsterdam/Philadelphia: John Benjamins.
- Altenberg, B. & Granger, S. (2002). *Lexis in Contrast. Corpus-based approaches*. Amsterdam/Philadelphia: John Benjamins Publishing Company
- Altenberg, B. & Granger, S. (2001). The Grammatical and Lexical Patterning of Make in Native and Non-native Student Writing. *Applied Linguistics* 22 (2), 173-195.
- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. En A. Cowie (Ed.), *Phraseology: Theory, analysis and applications*, 101-22. Oxford: Oxford University Press.
- Altenberg, B. & Eeg-Olofson, M. (1990). Phraseology in spoken English: presentation of a project. En Aarts, J. y Meijs, W. (Eds.), *Theory and Practice in Corpus Linguistics*, 1-26. Amsterdam: Rodopi.
- Álvarez Montalbán, F. & Albanesi, L. (2004). La enseñanza del español en Suecia. *El español en el mundo. Anuario del Instituto Cervantes*. Disponible en: http://cvc.cervantes.es/obref/anuario/anuario_04/alvarez_albanesi/default.htm
- Axelsson, M. & Westergren, Y. (2002). The Uppsala Student English Corpus (USE): A multi-faceted Resource for Research and Course Development. *Language and Computers: Studies in Practical Linguistics* 43. Amsterdam, Netherlands: Rodopi.
- Axelsson, M.W. (2000). USE - The Uppsala Student English Corpus: An instrument for needs analysis. *ICAME Journal*, 24, 155-157.

BIBLIOGRAFÍA

- Baker, M. (1995). Corpora in Translation Studies. An Overview and some suggestion for future research. *Target* 7(2), 223-243.
- Baralo, M. (2004). *La adquisición del español como lengua extranjera*. Madrid: ARCO/LIBROS, S.L.
- Baralo, M. (2010). La investigación en español como lengua segunda: necesidad de un corpus de español. Ponencia presentada en el V Congreso internacional de la lengua española. Disponible en:
<file:///G:/Ponencia%20de%20Marta%20Baralo%20IV%20Lengua%20y%20educaci%C3%B3n%20V%20Congreso%20Internacional%20de%20la%20Lengua%20Espa%C3%B1ola.htm>
- Bardel, C. (2000). *La negazione nell'italiano degli svedesi. Sequenze acquisizionali e influssi translinguistici*. Tesis doctoral. Lund: Universidad de Lund
- Barlow, M. (2005). Computer-based analysis of learner language. En R. Ellis and G. Barkhuizen (Eds.), *Analysing Learner Language*. Oxford: OUP.
- Beebe, L.M., Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. En R. C. Scarcella, E. S. Anderson, & S. D. Krashen (Eds.), *Developing communicative competence in a second language*, 55-94. New York: Newburry House.
- Benson, K. & Standvik, I. (1999). NORSTEDTS SVENSK-SPANSKA-SPANSK- SVENSKA ORDBOK. Stockholm: Norstedts Förlag.
- Berber Sardinha, T. (2000). Lingüística de Corpus: Histórico e Problemática. *D.E.L.T.A.*, Vol.16, No 2, 323-327. Disponible en:
<http://www.scielo.br/pdf/delta/v16n2/a05v16n2.pdf>
- Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, Vol. 8, No. 4, 1-15. Oxford: Oxford University Press.
- Biber, D. (2005). Paquetes léxicos en textos de estudio universitario: Variación entre disciplinas académicas. *Revista Signos*, 38 (57), 19-29. Disponible en:
http://www.scielo.cl/sciel0.php?pid=s071809342005000100002&script=sci_arttext 24-01-2008

- Biber, D. (2006). *University Language: A Corpus-Based Study of Spoken and Written Registers*. Philadelphia, PA, USA: John Benjamins Publishing Company.
- Bosque, I. & Demonte, V. (2000). *Gramática descriptiva de la lengua española*. Madrid: Editorial Espasa Calpe, S.A
- Braun, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, 19/3, 307-328.
- Butler, C. S. (1997). Repeated word combinations in spoken and written texts: some implication for Functional Grammar. En C.S. Buttler, J. H. Connolly, R. A. Gatward & R. M. Visman (Eds.), *A Fund of Ideas: Recent Developments in Functional Grammar*, 60-77. Amsterdam: IFOTT, University of Amsterdam.
- Butler, C. S. (2008). Three English adverbs and their formal equivalents in Romance languages, A Corpus-based collocation study. *Languages in Contrast* 8:1, 107–124. John Benjamins Publishing Company. doi: 10.1075/lic.8.1.10butissn 1387–6759 / e-issn 1569–9897.
- Butt, J. & Benjamin, C. (2000). *A New Reference Grammar of Modern Spanish*. London: ARNOLD.
- Carmona, J. et al. (1998). An environment for morphosyntactic processing of unrestricted Spanish text. *Proceedings of the First Conference on Language Resources and Evaluation*, 915–922. Granada: Universidad de Granada.
- Carter, R. (1988). *Vocabulary: Applied Linguistic Perspectives*. (2nd Ed.). London: Routledge.
- Cassany, D., Sanz, G. & Luna, M. (2007). Enseñar lengua. *Serie didáctica de la lengua y la literatura* 117. Barcelona: Editorial GRAÓ.
- Chambers, A. (2007). Popularizing Corpus consultation by language learners and teachers. En H. Hidalgo, L. Quereda & J. Santana (Eds.), *Corpora in the Foreign Language Classroom: selected papers from the Sixth International Conferencing on Teaching and Language Corpora*, 3-16. Amsterdam/Atlanta, Georgia: Rodopi.

BIBLIOGRAFÍA

- Carrió Pastor, M. & Seis Ortíz, R. (1999). La expresión escrita en inglés técnico. Sus errores. *Actes III Congrès Internacional sobre Llengües per a Finalitats Específiques*, 69-73. Barcelona: Publicacions de la Universitat de Barcelona.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System* 25 (3), 301-315.
- Connor, U. & Upton, T. A. (2004). *Discourse in the Professions. Perspectives from Corpus linguistics*. Amsterdam/London: John Benjamins Publishing Company.
- Cook, V. (2001). *Second Language Learning and Teaching*. London, New York: Arnold, Oxford University Press.
- Cook, V. (2002). Background to the L2 User. En V. Cook (Ed.), *Portraits of the L2 User*, 1-29. Clevedon: Multilingual Patters.
- Corder, S. P. (1967). The significance of learners' errors. En *IRAL* 5, 161-170, En S.P. Corder (1981), *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- Cortés Moreno, M. (2000). *Guía para el profesor de idiomas. Didáctica del español y segundas lenguas*. Barcelona: OCTAEDRO, S.L.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes* 23 (4), 397-423.
- Cresswell, J. (2003). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (2nd ed.). Thousand Oaks: Sage Publications.
- Cummin, A. (1989). Written expertise and second-language proficiency. *Language Learning* 39, 81-141.
- Danielsson, P. (2007). What Constitutes a Unit of Analysis in Language?. Birmingham, *Linguistic online* 31, 2/07, 1-8.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

- Davies, M. (2008). New directions in Spanish and Portuguese Corpus Linguistics. *Studies in Hispanic and Lusophone Linguistics*, Vol.1, Issue1, Spring 2008, 149-186.
- Davies, M. (2002-). Corpus del Español (100 million words, 1200s-1900s). Disponible en: <http://www.Corpusdelespanol.org>
- DeMello, G. (1998). Verbo de influencia + cláusula/infinitivo con sujetos no correferenciales. *Estudios en honor del profesor Josse de Kock*, 177-184. Leuven: Leuven University Press.
- Dewaele, J.M. (1998). Lexical inventions: French interlanguage as L2 versus L3. *Applied Linguistics* 19, 471-490.
- Domínguez García, N. (2007). *Conectores discursivos en textos argumentativos breves*. Madrid: Arco/Libros,S.L.
- Dulay, H. & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24, 37-53.
- Ellis, R. (2003). *Second Language Acquisition*. Oxford, New York: Oxford University Press.
- Ellis, R. (2007). Educational Settings and Second Language Learning. *Asian EFL Journal*, Vol. 9, No. 4, Conference Proceedings, 11-27.
- Enkvist, I. (2005). *Trängd mellan politik och pedagogik. Svenske språkutbildnings efter 1990*. Riga: Presses Nams. Disponible en:
http://gupea.ub.gu.se/bitstream/2077/18161/1/gupea_2077_18161_1.pdf
- Epstein, S. Flynn, S. & Martohardjono,G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioural and Brain Sciences* 19, 677-714. Copyright © Cambridge University Press 1996 DOI: 10.1017/S0140525X00043521. Published online: 04 February 2010.

BIBLIOGRAFÍA

- Erman, B. & Warren, B. (2000). The idiom principle and the open choice principle. *Text* 20(1), 29-62.
- Escandell Vidal, M.V. (2008). *Introducción a la pragmática*. Barcelona: Ariel lingüística.
- Estling, M. & Lindquist, H. (2007). Learning English grammar with a corpus: experimenting with concordancing in a university grammar course. *ReCall* 19 (3), 329-350. United Kingdom: Cambridge University Press. doi: 10.1017/S0958344007000638
- Ferm, R. & Malmberg, P. (2001). *Språkboken -en antologi om språkundervisning och språkinläring*. Örebro: Skolverket.
- Flowerdew, L. (1998). Integrating 'Expert' and 'Interlanguage' Computer Corpora Findings on Causality: Discoveries for Teachers and Students. *English for Specific Purposes* 17(4): 329-345.
- Flowerdew, L. (2000). Investigating referential and pragmatic errors in a learner corpus. En L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective*, 117–124. Frankfurt am main: Peter Lang Publishers.
- Flowerdew, L. (2002). Corpus-based analyses in EAP. En J. Flowerdew (Ed.), *Academic discourse*, 95-115. London: Longman.
- Flowerdew, L. (2003). A Combined Corpus and Systemic-Functional Analysis of the Problem-Solution Pattern in a Student and Professional Corpus of Technical Writing. *TESOL Quarterly* 37(3): 489-511.
- Flowerdew, L. (2004). The argument for using specialized corpora to understand academic and professional language. En U. Connor & T. A. Upton (Eds.), *Studies in Corpus Linguistics* 16, *Discourse in the Profession. Perspectives from corpus linguistics*, 11-33. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Flowerdew, L. (2008). *Corpus-based Analyses of the Problem- solution Pattern*. Amsterdam/London: John Benjamins Publishing Company.

- Flowerdew, L. (2009). Applying corpus linguistics to pedagogy. A critical evaluation. *International Journal of Corpus Linguistics* 14 (3), 393–417. doi: 10.1075/ijcl.14.3.05floissn 1384–6655 / e-issn 1569–9811 © John Benjamins Publishing Company
- Fält, Gunnar (2000): *Spansk Grammatik för universitet och högskolor*. Lund: Student litteratur.
- Forsberg, F. (2006). *Le langage préfabriqué en français parlé L2. Étude acquisitionnelle et comparative*. Tesis doctoral. Stockholm: Stockholm universitet.
- Forsberg, F (2010). Using conventional sequences in L2 French. *IRAL* 48, 25-51. DOI 10.1515/iral.2010.002
- Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. En M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing*, 75–93. Harlow, UK: Longman.
- Fuentes Rodríguez, C. & Alcaide Lara, E.R. (2007). *La argumentación lingüística y sus medios de expresión*. Madrid: Arco/Libros, S.L.
- Fulk, R. C. (1993). The Spanish of Mexico: A Partially Annotated Bibliography for 1970-90 Part I. *Hispania*, Volume 76, Number 2, May 1993.
Disponible en:
<http://www.cervantesvirtual.com/servlet/SirveObras/12059435340142617432435/p0000008.htm-010409>
- Gaskell, D. & Cobb T. (2004). Can learners use concordance feedback for writing errors?. *System* 32, 301-319.
- Gass, S. (1988). Second Language Acquisition and Linguistic Theory: The Role of Language Transfer. En S. Flynn and W. O'Neil (Eds.), *Linguistic Theory in Second Language Acquisition*, 384-403. Kluwer Academic Publishers.
- Gass, S. & Selinker, L. (1993). *Language transfer in Language Learning*. Rowley, Mass.: Newbury House.

BIBLIOGRAFÍA

- Gass, S. & Selinker, L. (2008). *Second Language Acquisition: An Introductory course* (3rd ed.). New York: Routledge.
- Gavioli, L. (2005). *Exploring Corpora for ESP Learning*. Amsterdam, NLD: John Benjamins Publishing Company. Disponible en: <http://site.ebrary.com/lib/jonhh/Doc?id=10103908&ppg=29>
- George, H. (1972). *Commons error in language learning*. Rowley, Mass.: Newbury House.
- Giacobbe, J. (1992). A cognitive view of the role of L1 in the L2 acquisition process. *Second Language Research*, 8, 232-50.
- Gilmore, A. (2008). Using online corpora to develop students' writing skills. *English Language Teaching Journal*, 63 (4), 363-372.
- Gilquin, G. (2005). *Linking up contrastive and learner corpus research: an introduction*. Paper presented at the workshop Linking up Contrastive and Learner Corpus Research, University of Santiago de Compostela, 19 September 2005.
- Gilquin, G. (2008). Combining contrastive and interlanguage analysis to apprehend transfer: detection, explanation, evaluation. En G. Gilquin, Papp S. & M. B. Díez-Bedmar (Eds.), *Linking up Contrastive and Learner Corpus Research*, 3-33. Amsterdam, Atlanta : Rodopi.
- Gilquin, G., Granger S. & Paquot M. (2007). Learner corpora: the missing link in EAP pedagogy. En P. Thompson (Ed.), *Corpus-Based EAP Pedagogy*. Special issue of *Journal of English for Academic Purposes* 6(4), 319-335.
- González Piñeiro, M. Guillén Díaz, C. & Vez, J.M. (2010). *Didáctica de las lenguas modernas, competencia plurilingüe e intercultural*. Madrid: EDITORIAL SÍNTESIS, S.A.
- Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual corpora and learner corpora. En Aijmer et al. (Eds), *Languages in Contrast*, 37-51.

- Granger, S. (1998). Prefabricated pattern in advanced EFL writing: collocations and formulae. En A.P.Cowie (Ed.). *Phraseology: Theory, Analysis and Applications*, 145-160. Oxford:OUP.
- Granger, S., Dagneaux, E. & Meunier, F. (2002). *International Corpus of Learner English*. Louvain: UCL Presses Universitaires de Louvain.
- Granger, S. (2003). A multi-contrastive approach to the use of link words by advanced learners of English: evidence from the International Corpus of Learner English. *Paper presented at the Pragmatic Markers in Contrast workshop organized by the Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, Brussels, 22-23 May 2003*.
- Granger, S. (2002). A Bird's-eye View of Computer Learner Corpus Research. En S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Language Learning and Language Teaching 6, 3-33. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Granger, S. (2005). Computer Learner Corpus Research: current status and future prospects. En U. Connor & T. Upton (Eds.), *Applied Corpus Linguistics: a multidimensional Perspective*, 123-145. Amsterdam & Atlanta: Rodopi.
- Granger, S. & M. Paquot (2005). The phraseology of EFL academic writing: Methodological issues and research findings. Paper presented at *ICAME 26 – AAAACL6* (International Computer Archive of Modern and Medieval English - American Association of Applied Corpus Linguistics), 12-15 May 2005, University of Michigan, USA.
- Granger, S. (2007). Integrating learner corpora and natural language processing: A crucial step towards reconciling technological sophistication and pedagogical effectiveness. *ReCall* 19, 3, 252-268.
- Hasselgren, A. (2002). Learner corpora and language testing: Small words as markers of learner fluency. En S. Grange, H. Joseph & S. Petch-Tyson

BIBLIOGRAFÍA

- (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, 143–173.
- Haugen, F. (1966). Dialect, language, nation. *American Antropologist*, 68, 922-35.
- Hawkin, R. (2001). The theoretical significance of Universal Grammar in second language acquisition. *Second Language Research* 17 (4), 345-367.
- Hayes, J. R. & Flower, L. S. (1980). Identifying the organization of writing processes. En L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach*, 3-30. Hillsdale, N.Y.: Lawrence Erlbaum Associates.
- Hayes, J.R. (1996). A New framework for understanding cognition and affect in writing. En C. M. Levy & S. Randell (Eds.), *The science of writing: Theories, methods, individual differences and applications*, 1-27. Mahwah, N. J.: Lawrence Erlbaum Associates.
- Herbst, T. (2009). Valency- item-specificity and idiom principle. En U. Römer & R. Schulze (Eds.), *Exploring the Lexis-Grammar Interface*, 48-68. John Benjamins Publishing Company: Amsterdam/Philadelphia
- Herwig, A. (2001). Plurilingual lexical organisation: Evidence from lexical processing in L1-L2-L3-L4 translation. En J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistics influence in Third Language Acquisition: Psycholinguistic Perspectives*, 115-137. Clevedon: Multilingual Matters.
- Hejzlar, J. (2005). Utvärderingar av grund- och forskarutbildning i romanska språk vid svenska universitet och högskolor, Del II: spanska. *Högskoleverket rapportserie 2005:3r*.
- Helms-Park, R. (2001). Evidence of lexical transfer in learner syntax: the acquisition of English causatives by speaker of Hindi-Urdu and Vietnamese. *Studies in Second Languages Acquisition*, 23, 71-102.
- Hoey, M. (2009). Corpus-driven approaches to grammar. En U. Römer & R. Schulze (Eds.), *Exploring the Lexis-Grammar Interface*, 32-47. Benjamins Publishing Company: Amsterdam/Philadelphia.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

- Holm, B. & Nylund, E. (1979). *Deskriptiv svensk grammatik*. Trelleborg: Språkförlaget Skriptor AB.
- Hunston, S. (2006). *Corpora in Applied linguistics*. (4ta ed.). UK: Cambridge University press.
- Jarraud, A., Ramón, S., Jarraud, F. & Ramón, H. (2008). *Diccionario de uso del español*. (Edición abreviada). Madrid: Editorial Gredos, S.A.U.
- Jarvis, S. (1998). *Conceptual Transfer in the Interlanguage lexicon*. Bloomington: Indiana University Linguistics Club.
- Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, 50, 245-309.
- Johansson, S. (2007). *Seeing Through Multilingual Corpora: On the Use of Corpora in Contrastive Studies*. Amsterdam: John Benjamin Publishing Company.
Disponible en:
<http://site.ebrary.com.bibl.proxy.hj.se/lib/jonhh/docDetail.action?docID=10161067>
- Johansson, S. (2009). Some thoughts on corpora and second-language acquisition. En K. Aijmer (Ed.), *Corpora and Language Teaching*, 33–44.
- Kaltenböck, G. & Mehlmauer-Larcher, B. (2005). Computer corpora and the language classroom: On the potential and limitations of computer corpora in language teaching. *ReCALL*, 17, 65-84.
- Kellerman, E. & S. Smith (1986). *Crosslinguistic influence in Second Language Acquisition*. Oxford: Pergamon Press.
- Kern, R. (2006). Perspectives on technology in learning and teaching languages. *TESOL Quarterly*, 40 (1), 183-210.
- Koike, N. (2007). *Spanish Learner Corpus and Exercises (SLCE)* de la Universidad de Texas. Disponible en <http://www.laits.utexas.edu/slce>

BIBLIOGRAFÍA

- Kopotev, M. V. (2003). CORPUS LINGUISTICS IN FINLAND: A RESOURCE SURVEY. *Automatic Documentation and Mathematical Linguistics*, Vol. 37, 3, 35-42.
- Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistic for Language Teachers*. Ann Arbor, Michigan: University of Michigan Press
- Larsen-Freeman, D. & M.H. Long (1993). *An Introduction to Second Language Acquisition Research*. London: Longman.
- Laufer, B. (2000). Avoidance of idioms in a second language: The effect of L1-L2 degree of similarity. *Studia Linguistica*, 54 (2), 186-196.
- Lazaraton, A. (2000). Currents trends in research methodology and statistics in applied linguistics. *TESOL Quaterly*, 34, 175-181.
- Lazaraton, A. (2003). Evaluative criteria for qualitative research in applied linguistics: Whose criteria and whose research?. *Modern Language Journal*, 87, 1-12.
- Leech, G. (1997). Teaching and language corpora: a convergence. En A. Wichmann, S. Fligestone, T. McEnery and G. Knowles (Eds.), *Teaching and Language Corpora*. Harlow: Addison Wesley Longman.
- Lindberg, J. (1997). *Korpuslingvistik. Del 1*. Disponible en: <http://www.ling.su.se/staff/beb/korpus.htm>
- Lope Blanch, J. M. (1986). *El estudio del español hablado culto: Historia de un proyecto*. México: UNAM.
- López S., A. & Borreguero Z., A.M. (2010). Los marcadores del discurso y la variación lengua hablada vs lengua escrita. En Loreda Lamas, Ó. & E. Acín Villa (coords), *Los estudios sobre marcadores del discurso en español, hoy* (pp. 415-495). Madrid: ARCO/LIBROS, S.L.
- Loureda Lamas, Ó. & Acín Villa, E. (2010). *Los estudios sobre marcadores del discurso en español, hoy*. Madrid: Arco/Libros, S.L.

- Loureda Lamas, Ó. (2009). *Introducción a la tipología textual*. Madrid: Arco/Libros, S.L.
- Lozano, C. (2008). CEDEL2. Corpus Escrito del Español L2/L2 Spanish written corpus. *International Seminar on Corpus Linguistics, New Trends in Language Teaching and Translation Studies*. Granada: Universidad de Granada.
- Lozano, C. & Mendikoetxea, A. (2007). Learner corpora and the acquisition of word order: A study of the production of verb-subject structures in L2 English. En M. Davies, Prayson, S. Hunston, & P. Danielsson (Eds.), *Proceedings of the Corpus Linguistics Conference 2007*. Birmingham: University of Birmingham.
- Lüdeling, A. Kytøe, M. & McEnery, T. (2008). *Corpus Linguistics: An International Handbook*. Berlin: Mouton de Gruyter.
- Lozano González, L (2005). Hacia una única explicación del subjuntivo aplicado a la adquisición de E/LE (2da parte: alternancia indicativo-subjuntivo). *Cuadernos Cervantes de la lengua española*, Año No. 11, No. 57. Disponible en:
http://www.cuadernos cervantes.com/ele_56_subjuntivo.htm
- Llisterri, J. & Garrido, J.M. (1998). La ingeniería Lingüística en España. *El español en el mundo. Anuario del Instituto Cervantes 1998*. Disponible en:
http://cvc.cervantes.es/obref/anuario/anuario_98/llisterri
- Llisterri, J. (2004). Las tecnologías lingüísticas en España. *El español en el mundo. Anuario del Instituto Cervantes 2004*. Disponible en:
http://cvc.cervantes.es/obref/anuario/anuario_04/llisterri/p04.ht
- Maqueo, A.M. (2005). *Redacción*. México: Limusa.
- Marín, F. M. & Sánchez Lobato, J. (1991). *Lingüística Aplicada*. Madrid: EDITORIAL SÍNTESIS, S. A..
- Martí, M. & Torrens, M.J. (2001). *Construcción e interpretación de oraciones. Los conectores oracionales*. Madrid: Edinumen.

BIBLIOGRAFÍA

- Martí Sánchez, M. (2008). *Los marcadores en español L/E: conectores discursivos y operadores pragmáticos*. Madrid: ARCO/LIBROS, S.L.
- Martí Sánchez, M.; Penadés Martínez, I. & Ruiz Martínez, A. M. (2008). *Gramática Española por niveles*, Volumen 2. Madrid: Edinumen.
- Martínez Sánchez, R. (2004a). La escritura en lengua extranjera: procesos y resultados. *La enseñanza de lenguas extranjeras desde una perspectiva europea*, 87-114). Madrid: Ministerio de educación y ciencia.
- Martínez Sánchez, R. (2004b). Manos a la letr@: la destreza escrita en el aula. En F. Meno Blanco (Ed), *Aprendizaje de las lenguas extranjeras en el marco europeo*, 87-96. Madrid: Ministerio de educación y ciencia.
- McCarthy, M. J. (2001). What is an advanced level vocabulary?. *Studies in English Language and Linguistics* 3, 149-163.
- McEnery, T. (2003). Corpus linguistics. *The Oxford Handbook of Computational Linguistics* , 448-463.
- McEnery, T.; Xiao, R. & Tono, Y. (2006). *Corpus-based Language Studies: An Advanced Resource Book*. Great Britain: Routledge Taylor & Francis Group Plc.
- McEnery, T. & Wilson, A. (2001). *Corpus Linguistics: an Introduction*. Edinburgh: EDINBURGH University Press.
- McEnery, T. & Wilson A. (1996). *Corpus Linguistics*. Edinburgh: EDINBURGH University Press.
- Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching. En S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, 119–142. Philadelphia: John Benjamins.
- Miche, E. (1998). Secuencias discursivas del desacuerdo. *Verba: Aplicación del modelo ginebrino al análisis del debate parlamentario del artículo 2 de la Constitución*

española de 1978. Santiago de Compostela: Servicio de Publicacións e Intercambio científico.

Mighetto, D. (1998-). Corpus SOL- Spanish Online. Concordancias españolas en la red. Disponible en: <http://spraakbanken.gu.se/konk/rom2/>

Mitchell, T., Shinkareva, S., Carlson, A., Chang, K., Malave, V., Mason, R. & M. Just. (2008). Predicting human brain activity associated with the meanings of nouns. *Science*, 320, 1191–1195. DOI:10.1126/science.1152876

Monné Marsellés, P. (1998). La escritura y su aprendizaje. En A. Mendoza Fillola (Coordinador). *Conceptos claves de didáctica de la lengua y la literatura*. Barcelona: SEDLL, ICE Universitat de Barcelona, Edit. Horsori.

Montolío, E. (2001). *Conectores de la Lengua Escrita*. Barcelona: Ariel Practicum.

Montrul, S. (2000). Transitivity alternations in L2 acquisition. Toward a modular view of transfer. *Studies in Second Language Acquisition* 22/2, 229-233.

Moon, R. (2007). Sinclair, lexicography, and the Cobuild Project. *International Journal of Corpus linguistics* 12 (2), 159-181.

Moon, R. (1997). Vocabulary connections: multi-word items in English. En N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy*, 40-63. Cambridge: CUP.

Moreno Teva, I. (2007). Estudio evolutivo de las secuencias formulaicas en aprendientes suecos de E/LE. *Actas del II Congreso de Hispanistas y Lusitanistas Nórdicos*. Stockholm: Universidad de Estocolmo.

Moskowich-Spiegel Fandiño, I., Crespo García, B., Lareo Martín, I. & Lojo Sandino, P. (2010). *Language Windowing through Corpora. / Visualización del lenguaje a través de corpus*. II International Conference on Corpus Linguistics. A Coruña: Universidade da Coruña.

BIBLIOGRAFÍA

- Nesselhauf, N. (2004) Learner corpora and their potential for language teaching. En Jonh McHardy Sinclair (Ed.), *How to use corpora in language teaching*. Amsterdam: Jonh Benjamins Publishing Company.
- Nesselhauf, N. (2005): *Collocation in a Learner Corpus*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- Odling, T. (1989). *Language Transfer*. Cambridge: Cambridge University Press.
- Odling, T. (2003). Cross-linguistic influence. En C. Doughty & M. Long (Eds.), *Handbook of second language acquisition*, 436-486. Oxford: Blackwell.
- Österberg, R. (2008). *Motivación, aptitud y desarrollo estructural: un estudio sobre la actuación lingüística en aprendientes suecos de español L2*. Tesis doctoral. Stockholm: Stockholms universitet.
- Paquot, M. (2005). Towards a productively-oriented academic word list. Paper presented at *Practical Applications in Language and Computers*, 7-9 April 2005, Łódź, Poland.
- Parodi, G. (2010). *Lingüística de corpus: de la teoría a la práctica*. Madrid: Iberoamericana.
- Pastor Cesteros, S. (2005). La enseñanza de segunda lengua. En A. López & B. Gallardo (Eds.), *Conocimiento y lenguaje*, 361-399. Valencia: Publicaciones de la Universidad de Valencia.
- Pawley, A. & Syder, F. (1983). *Two Puzzles for linguistic theory: nativelike selection and nativelike fluency*. En J. Richards & R. Schmidt (Eds.), *Language and Communication*. Longman.
- Penny, R. (1993). *Gramática histórica del español*. Barcelona: Editorial Ariel, S.A.
- Pérez Hernández, C. (2002). Explotación de los córpora textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento. Disponible en: <http://elies.rediris.es/>
- Pino, A. (2006). Variaciones en contexto: un análisis del discurso académico escrito por suecos aprendices de español como lengua extranjera. *Trabajo*

de DEA, programa de doctorado de lingüística aplicada, curso 2006/07.

Universitat de Barcelona

- Pino, A. (2010). Que en contexto: un estudio del uso y variación de combinaciones de palabras por aprendices suecos de E/LE. En J. L. Bueno Alonso, D. González Álvarez, U. Kirsten Torrado, A. E. Martínez Inzúa, J. Pérez Guerra, E. Rama Martínez & R. Rodríguez Vázquez (Eds.), *Analitzar dades > Describir variación / Analysing data > Describing variation*, 92-100. Vigo: Universidade de Vigo (Servizo de Publicacións).
- Porquier, R. (1991). Pour une conceptualisation contrastive au niveau avancé. En R. Gauchola, C. Mestreit & M. Tost. (coord.), *Méthodologie, formation, pragmatique et analyse textuelle. Experiences et pratiques d'enseignement-aprentissage de FLE. Actas de las XIVas Jornadas pedagógicas sobre la enseñanza del francés en España*. Barcelona: ICE de la Universidad Autónoma de Barcelona (pp. 11-20).
- Postigo, R. M. (1998). Modelos de educación lingüística en contextos de lenguas en contacto. En A. Mendoza Fillola (coord.), *Conceptos claves en Didáctica de la Lengua y la Literatura*. SEDLL/ ICE/ HORSORI, 265-274.
- Quirk, R. Greenbaum, S. Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Rayson P., Berridge D. and Francis B. (2004). Extending the Cochran rule for the comparison of word frequencies between corpora. En G. Purnelle, C. Fairon & A. Dister (Eds.), *Le poids des mots: Proceedings of the 7th International Conference on Statistical analysis of textual data (IADT 2004)*, 926 – 936. Louvain-la-Neuve, Belgium, March 10-12, 2004, Presses universitaires de Louvain.
- Rayson, P. and Garside, R. (2000). Comparing corpora using frequency profiling. En proceedings of the workshop on *Comparing Corpora, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000)*. 1-8 October 2000, Hong Kong, 1-6.

BIBLIOGRAFÍA

- Real Academia Española/Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española. Sintaxis II*. Madrid: Espasa libros, S.L.U.
- Richards, J. C. (1971). Error analysis and second language strategies. *Language Science*, 17, 12-22.
- Ringbom, H. (2007). *Cross-linguistic Similarity in Foreign Language Learning*. Clevedon, UK: Multilingual Matters.
- Römer, U. (2010). Establishing the phraseological profile of a text type: The construction of meaning in academic book reviews. *English Text Construction* 3(1), 95-119.
- Römer, U. & Schulze, R. (2009). *Exploring the Lexis- Grammar Interface*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sánchez, A. (2001). Investigación y análisis mediante corpus lingüísticos: el poder de atracción de las palabras. En P. Fernández Nistal y J.M. Bravo Gonzalo (Eds.), *Pathways of translation studies*, 11-46. Valladolid: Universidad de Valladolid.
- Sánchez, A. (1995). *Cumbre, corpus lingüístico del español contemporáneo*. Madrid: SGEL- Educación.
- Sánchez, A. & Cantos, P. (2009). *A Survey on Corpus-based Researchs7 Panorama de investigaciones basadas en corpus*. Murcia: Editum (Servicio de publicaciones de la Universidad de Murcia).
- Sánchez-Reyes, S. (2004). El texto, vehículo de comunicación entre individuos y entre pueblos. En F. Meno Blanco (Ed.), *La enseñanza de las lenguas extranjeras desde una perspectiva europea*, 63-86. Madrid: Ministerio de educación y ciencia.
- Sánchez Quintana, N. & CLARK, D. (1995). *Destrezas integradas. Las cuatro destrezas*. Madrid: SM.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

- Santos Gargallo, I. (1993). Análisis contrastivo, Análisis de errores e Interlengua en el marco de la lingüística contrastiva. Madrid: Síntesis.
- Scott, M. *WordSmith Tools. Versión 5*. Licencia, Universidad de Gotemburgo. Suecia. <http://www.lexically.net/wordsmith/>
- Schmitt, N. & Carter, R. (2004). Formulaic sequences in action: An introduction. En N. Schmitt (Ed.), *Formulaic Sequences*, 1-22. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Scardamalia, M. & Bereiter, C. (1992). Dos modelos explicativos de los procesos de comprensión escrita. *Infancia y aprendizaje*, 58, 43-64.
- Selinker, L. (1969). Language transfer. *General Linguistics*, 9, 67-92.
- Selinker, L. (1972). Interlanguage. *IRAL*, X (2), 209-231.
- Seidlhofer, B. (2009). Accommodation and the idiom principle in English as a Lingua Franca. *Intercultural Pragmatics* 6-2, 195-215. DOI 10.1515/IPRG.2009.011
- Siepmann D. (2005). *Discourse Markers across Languages. A contrastive study of second-level discourse markers in native and non-native text with implications for general and pedagogic lexicography*. London: Routledge.
- Siguán, M. & Mackey, W. F. (1986). *Educación y bilingüismo*. Madrid: UNESCO/Santillana.
- Simpson, R. (2004). Stylistic features of academic speech: The role of formulaic expressions. En U. Connor & T. A. Upton (Eds.), *Discourse in the Professions*, 37-64.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford/ New York, Oxford University Press.
- Sinclair, J. (1996). The Search for Units of Meaning. *Textus IX*, 75-106.

BIBLIOGRAFÍA

- Sinclair, J. (2005). Corpus and text- Basic principles. En M. Wynne (Ed.), *Developing Linguistics Corpora: A guide to good practice*. Oxford: Oxbow Books.
- Siyanova, A & Schmitt, R. (2008). L2 Learner Production and Processing of Collocation: A Multi-study Perspective. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 64, 3 (March/mars), 429–458.
- Säljö, R. (2000). *Lärande i praktiken, ett sociokulturellt perspektiv*. Stockholm: Prisma.
- Söhrman, I. (2007). *La lingüística contrastiva como herramienta para la enseñanza de lenguas*. Madrid: Arco/Libros, S.L.
- Stanley Whitley, M. & González, L. (2007). *Gramática para la composición* (2da Ed.). Washington D.C: Georgetown University Press.
- Stubbs, M. (1996). *Text and Corpus Linguistics*. Oxford: Blackwell
- Tagnin, S. (2002) *USP Multilingual Learner Corpus* (MLC). Departamento de Lenguas modernas de la Universidad de San Pablo (Brasil)
- Tarone, E. (1979). Interlanguage as chameleon. *Language Learning* 29, 181-191.
- Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning*, 32, 69-82.
- Teleman, U., Hellberg, S. & Andersson, E. (1999). *Svenska Akademiens grammatik 2 Ord*. Stockholm: Svenska Akademien.
- Teubert, W. & Cermáková, A. (2007). *Corpus Linguistics. A Short Introduction*. London: Atheneum Press.
- Tizón Couto, B. (2009). Un estudio basado en corpus de la complementación clausal dependiente de verbos en el inglés oral de aprendices y nativos: comparando VICOLSE, LINDSEI Spa y LOCNEC. En R. Caballero Rodríguez & M. Jesús Pinar Sanz (Eds.), *Modos y formas de la comunicación humana*, XXVII Congreso Internacional de AESLA. Cuenca: UCLM.

EL USO DE COMBINACIONES DE PALABRAS CON *QUE* EN UN CORPUS DE APRENDICES
SUECOS DE ESPAÑOL COMO LENGUA EXTRANJERA

- Tognini-Bonelli, E. (2002). *Corpus Linguistic at Work*. Amsterdam/ Philadelphia: John Benjamins.
- Torruella, J. & Llisterri, J. (1999). Diseño de corpus textuales y orales. En BLECUA, J.M., CLAVERÍA, G., SÁNCHEZ, C. & TORRUELLA, J. (Eds.), *Filología e informática. Nuevas tecnologías en los estudios filológicos*, 45-77. Barcelona: Universidad Autónoma de Barcelona, Editorial Milenio.
- Verdaguer Clavera, I. (2004). Word Meaning and Collocations in Scientific English. *Revista de Inglés para Fines Específicos*, 1, 93-97.
- Verdaguer Clavera, I. (2003). Collocation in Scientific Language: a contrastive study. *Studies in Contrastive Linguistics*. Proceedings of the 3rd International Contrastive Linguistics Conference, 633-639.
- Weinreich, U. (1953). *Languages in Contact*. The Hague: Mouton.
- Widdowson, H. G. (1991). The description and prescription of language. En J. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics*, 1991, 11-24. Washington, D.C.: Georgetown University Press.
- Widdowson, H. G. (1998). Context, community and authentic language. *TESOL Quarterly*, 32 (4), 705-716.
- Wiktorsson, M. (2003). *Learning idiomaticity: A corpus-based study of idiomatic expressions in learners' written production*. Lund Studies in English 105. Tesis doctoral. Lund: Departamento de Inglés, Universidad de Lund.
- Wijk-Andersson, E. (1981). *Ny Grammatik. Det svenska språkets struktur*. Uppsala: Hallgren & Fallgren.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: University Press.
- Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching*, 32, 213-231.

BIBLIOGRAFÍA

Otras fuentes electrónicas consultadas:

Asociación española de lingüística de corpus. Disponible en:

<http://www.um.es/aelinco/>

Avhandlingar.se. Disponible en: <http://www.avhandlingar.se/>

Banco de tecnología lingüística. Disponible en: <http://sprakteknologi.se/>

CEDEL2. Corpus escrito del español L2. Disponible en:

<http://www.uam.es/woslac/cedel2.htm>

Center for English Corpus Linguistic, Universidad católica de Louvain. Disponible

en: <http://www.uclouvain.be/en-cecl-icle.html>

Center for Corpus Research. Universidad de Birmingham. Disponible en:

<http://www.birmingham.ac.uk/research/activity/corpus/index.aspx>

Corpus del español. Disponible en: <http://www.corpusdelespanol.org/>

Corpus Sol Spanish- Online. Disponible en:

<http://spraakbanken.gu.se/swe/resurs/cor92>

CrossCheck, a grammar checking system especially developed for second language learners of

Swedish. Disponible en:

<http://www.csc.kth.se/tcs/projects/xcheck/index.html> UCRE

Database. ESF (European Science Foundation Second Language). Disponible en:

<http://www.mpi.nl/world/tg/lapp/esf/esf.html>

Diccionario de la real academia española. Disponible en: <http://www.rae.es/>

Diccionario de terminos claves de ELE. Disponible en:

http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/default.htm

[m](#)

Dirección de educación superior sueca. Disponible en: <http://www.hsv.se/>

Informes de la Dirección de educación superior sueca. Disponibles en:

<http://www.hsv.se/download/18.539a949110f3d5914ec800090615/isbn91-88874-71-0.pdf>- 2009-12-05

<http://www.hsv.se/download/18.539a949110f3d5914ec800092239/0503R-2.pdf#search=%27spanska%27->

<http://www.hsv.se/download/18.1dfeedbc1284e5ecc3d8000350/1006R-uppfoljning-utvarderingar-2004.pdf>

<http://www.hsv.se/densvenskahogskolan/universitetochhogskolor/adresser.4.539a949110f3d5914ec800062390.html>

<http://www.hsv.se/download/18.ac0b56a12a53b8af8580001119/1010R-universitet-hogskolor-arsrapport-2010-utbild-grund-avancerad.pdf>

El Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación. 2001. Estrasburgo: Consejo de Europa. Departamento de política lingüística. Centro Virtual Cervantes. Disponible en:
http://cvc.cervantes.es/obref/marco/cap_03_01.htm.

Proyecto PRESEEA. Disponible en:

<http://www.linguas.net/Default.aspx?alias=www.linguas.net/portalpreseea>

Project SPRIK. Språk i kontrast. Disponible en:

<http://www.hf.uio.no/ilos/forskning/prosjekter/sprik/>

Språkkontakt og ungdomsspråk i Norden. Disponible en: <http://www.uib.no/uno/>

The European Language Portfolio (ELP). Disponible en:

http://www.coe.int/t/dg4/portfolio/default.asp?l=e&m=/main_pages/welcome.html

The Helsinki Corpus of English Texts. Disponible en:

<http://khnt.hit.uib.no/icame/manuals/hc/INDEX>.

University Centre for Computer Corpus Research on Language. Universidad de Lancaster. Disponible en: <http://ucrel.lancs.ac.uk>

LISTA DE ANEXOS

1. Plan de estudio del *Grundkurs 1-30hp*, cursos 2008/09 y 2009/10. Facultad de humanidades. Universidad de Gotemburgo.
2. Plan de estudio del curso *Spanska 1-30hp*, cursos 2008/09 y 2009/10. Facultad de educación y comunicación. Instituto superior de Jönköping.
3. Carta de presentación de la tesis a los participantes.
4. Certificado de consentimiento informado.
5. Encuesta.



GÖTEBORGS UNIVERSITET
Humanistiska fakultetsnämnden

SP1101, Grundkurs I, 30 högskolepoäng
Spanish: Introductory Course I, 30 HECR
Grundnivå

1. Fastställande

Kursplanen är fastställd av Humanistiska fakultetsnämnden 2006-11-20 att gälla from ht 2007. Reviderad version gäller from vt 2008.

2. Inplacering

Kursen ges som fristående kurs.

3. Förkunskapskrav

Utöver grundläggande behörighet fordras svenska/svenska som andraspråk 3 årskurser eller kurs B samt spanska steg 3 eller spanska B-språk 2 årskurser, alternativt C-språk 3 årskurser eller B-språk kurs B, alternativt C-språk kurs B eller motsvarande (betyget i vart och ett av dessa ämnen skall vara lägst 3), alternativt Praktisk spanska (SP1001 enligt de nya kursplanerna eller SP0500 eller SP0600 enligt de tidigare kursplanerna) eller Orienteringskurs i spanska, allmän inriktning, 20 poäng (SP1000) med godkänt resultat.

4. Innehåll

Kursen består av två delkurser:

4.1 Språkfärdighet och grammatik (15 hp)

Proficiency and grammar (15 HECR)

Delkursen innehåller praktiska övningar i muntlig och skriftlig framställning samt läs- och hörförståelse som examineras genom ett slutprov i två delar, i det ena bedöms den muntliga språkfärdigheten och hörförståelsen och i den andra den skriftliga språkfärdigheten.

Språkfärdighet, 5 hp

Delkursen behandlar även grunderna i allmän och spansk grammatik samt ger en introduktion i spansk fonologi. I kursen betonas de typologiska skillnaderna mellan svensk och spansk grammatik som kan innebära särskilda svårigheter för inlärare. Dessa delar av kursen examineras genom fyra delprov:

Allmän grammatik, 2 hp
Verbprov, 1 hp
Spansk Fonologi, 2 hp
Spansk Grammatik, 5 hp

4.2 Text, kultur och samhälle (15 hp)

Text, Culture and Society (15 HECR)

Huvudinnehållet är studium och diskussion av nutida spanskspråkiga texter. Delkursen behandlar kultur- och samhällsförhållanden i anslutning till de texter som läses och tar därvid upp t.ex. förekomsten av kulturella skillnader vid texttolkning, genusperspektiv samt miljö- och etnicitetsfrågor. Kursen examineras kontinuerligt genom inlämningsuppgifter och muntliga presentationer samt genom två prov, ett i mitten av terminen (5 hp) och ett slutligt prov (10 hp).

5. Mål

Studenten ska efter avslutad kurs:

- ha uppnått nivån B1 i hörförståelse, textförståelse, samtal/muntlig interaktion, muntlig produktion och skriftlig färdighet i spanska enligt den europeiska referensramen för språk,
- kunna använda en grundläggande grammatisk begreppsapparat,
- kunna förstå och tillämpa det spanska ljudsystemet och vara förtrogen med skillnaderna mellan svenskt och spanskt uttal,
- tillämpa grundläggande ordbildningslära i sin språkproduktion,
- behärska de regelbundna och oregelbundna formerna, i synnerhet inom verbläran,
- kunna använda sig av spansk syntax, framförallt kunna lösa konkordansproblem,
- behärska och kunna tillämpa kunskapen om de typologiska skillnaderna mellan svensk och spansk grammatik som komplicerar språkfärdigheten i spanska,
- kunna analysera olika texttyper, som t ex. skönlitteratur, film, tidningsartiklar,
- kunna formulera frågor utifrån texterna på begriplig spanska,
- kunna tillämpa några enklare teoretiska begrepp på texter vid analys av olika genrer,
- kunna diskutera kulturella likheter och skillnader med utgångspunkt från texterna,

6. Kurslitteratur

Se separat litteraturlista.

7. Former för bedömning

För godkänd kurs är deltagande i gruppövningar som förekommer under delkursen Språkfärdighet och grammatik är obligatoriskt.

Student äger rätt till byte av examiner efter att ha underkänts två gånger på samma examination, om det är praktiskt möjligt. En sådan begäran ställs till institutionen och skall vara skriftlig.

För studerande som ej blivit godkänd vid ordinarie prov ges ytterligare ett provtillfälle i regel inom fyra veckor efter ordinarie prov.

Den som godkänts i prov får ej undergå förnyat prov för högre betyg.

8. Betyg

Betygsskalan omfattar betygsgraderna Underkänd (U), Godkänd (G) och Väl godkänd (VG). Annan betygsskala förutsätter beviljad dispens.

För sammanfattningsbetyget Väl godkänd krävs detta vitsord på minst halva kursen.

Möjlighet att överföra betyget till ECTS-skalan om studenten i början av terminen ansöker om detta.

9. Kursvärdering

Varje delkurs ska utvärderas och resultaten ska bli föremål för diskussion mellan lärarna på kursen och representanter för studenterna. Protokoll/minnesanteckningar från denna diskussion skall avrapporteras till avdelningsnämnden. Utvärderingen skall ske dels efter avslutad kurs och dels vid minst ett tillfälle under kursens gång. Vid planering av påföljande kurstillfälle skall det dokumenteras hur resultaten av utvärderingen har tagits till vara.

Kursansvarig lärare är ansvarig för att studenternas synpunkter på kursen systematiskt och regelbundet inhämtas. Studierektor ansvarar för att resultaten av utvärderingar i olika former tas i beaktande vid revisioner av kursplanen.

10. Övrigt

Undervisningsspråket är huvudsakligen spanska.

Kursplanen fastställd av VD 2007-06-18

Gäller ht-08

Reviderad 2008-06-23

Kurskod: LSAA17

Ämnesnivå: A

Utb.omr.: HU

Ämneskod: SPA

Spanska, 1-30 högskolepoäng

Spanish, 1-30 credit points

Grundnivå

Allmänt

Kursen är en fristående kurs som syftar till att ge en allsidig praktisk och teorisk utbildning i spanska som modernt språk.

Lärandemål

Efter avslutad kurs förväntas de studerande ha

- utvecklat förmågan att använda spanska i tal och skrift utan att göra elementära uttals-, stavnings-, och grammatikfel.
- utvecklat kunskaper och färdigheter i grammatiska och fonetiska grundbegrepp och relevant terminologi samt tillägnat sig insikt i systematiska skillnader mellan svenska och spanska.
- utvecklat insikt och kunskap om företrädesvis modern spanskspråkig litteratur genom studier av texter (skönlitteratur, sakprosa) och olika genrer.
- utvecklat kunskaper om och skaffat sig överblick över de spanskspråkiga ländernas kultur, historia och samhällsförhållanden.

Innehåll

1. Muntlig språkfärdighet och fonetik, 7,5 högskolepoäng
2. Skriftlig språkfärdighet och grammatik, 7,5 högskolepoäng
3. Realia och sakprosa I, 7,5 högskolepoäng
4. Modern spanskspråkig litteratur, 7,5 högskolepoäng

Delkurs 1: Muntlig språkfärdighet och fonetik, 7,5 hp

Phonetics and Oral Proficiency, 7,5 credit points

Innehåll och former

Studium i spansk fonetik och fonologi med träning i tal och samtalsövningar.

En översikt över spanska språkets olika varianter samt stilnivåer både i Spanien och Latinamerika.

Undervisningen sker i form av föreläsningar, seminarier, gruppövningar, muntlig filmanalys och uttalsövningar.

Lärandemål

Efter avslutad delkurs förväntas de studerande

- känna till och förklara grundläggande skillnader mellan svenska och spanska vad gäller språkljudens och prosodins funktion
- kunna tillämpa teoretiska begrepp i fonetik och fonologi i den egna kommunikationsprocessen på målspråket
- förstå olika varianter av enklare spanskt tal
- kunna uttrycka sig muntligt med ett gott ordförråd

Examination

Delkursen examineras genom en skriftlig tentamen (salstentamen) i fonetik och muntliga examinatorier (uttal och muntlig framställning). För Godkänd på delkursen krävs att alla delar är fullgjorda och godkända. För Väl Godkänd krävs dessutom att de individuella proven har sammanräkningsbetyget Väl Godkänd.

Delkurs 2: Skriftlig språkfärdighet och grammatik, 7,5 hp

Grammar and Written Proficiency, 7,5 credit points

Innehåll och former

Studium i spansk ordbildning, verblära och allmän grammatik.

Träning i spansk skriftlig språkfärdighet.

Översättningsövningar.

Resuméer, skriftliga textredovisningar av skönlitteratur.

Undervisningen sker i form av föreläsningar, gruppövningar och individuellt arbete med övningsuppgifter.

Lärandemål

Efter avslutad delkurs förväntas de studerande kunna

- skriva på spanska utan att bryta mot elementära grammatiska regler
- översätta enklare texter
- tillämpa allmänna grammatiska grundbegrepp, behärska grammatisk terminologi samt kunna böja spanska verb
- förklara språkliga skillnader och likheter mellan svenska och spanska

Examination

Delkursen examineras genom en skriftlig tentamen (salstentamen) i grammatik och översättning samt två individuella inlämningsuppgifter. För betyget Godkänd på delkursen krävs att alla delar är fullgjorda och godkända. För Väl Godkänd krävs dessutom att de individuella proven har sammanräkningsbetyget Väl Godkänd.

Delkurs 3: Realia och sakprosa, 7,5 hp

Cultural and not Literary Texts, 7,5 credit points

Innehåll och former

Studium i Spaniens och Latinamerikas viktigaste geografiska, historiska, kulturella och sociala särdrag.

Analys av aktuella ekonomiska, politiska och sociala förhållanden ur ett historiskt perspektiv.

Redovisning av olika tidningsartiklar.

Undervisningen sker i form av föreläsningar, seminarier, gruppövningar och muntliga presentationer.

Lärandemål

Efter avslutad delkurs förväntas de studerande kunna

- uttrycka sig om den aktuella situationen i Spanien och Latinamerika med utgångspunkt i ekonomiska, politiska och historiska förhållanden
- söka, värdera och diskutera information om Spanien och de latinamerikanska länderna
- analysera sakprosatexter och presentera dem både muntligt och skriftligt

Examination

Delkursen examineras genom en muntlig presentation, (individuellt och i grupp) två inlämningsuppgifter (resuméer av två aktuella tidningsartiklar) och aktivt deltagande i seminarierna. För betyget Godkänd på delkursen krävs att alla delar är fullgjorda och godkända. För Väl Godkänd krävs dessutom att den individuella presentationen och minst en av de två inlämningsuppgifterna har betyget Väl Godkänd.

Delkurs 4: Modern spanskspråkig litteratur, 7,5 hp

Modern Spanish Language Literature, 7,5 credit points

Innehåll och former

Läsning och diskussion om modern spanskspråkig litteratur från 1900-talet till idag.

Extensiv och intensiv läsning av valda skönlitterära texter, d.v.s. studium av textens innehåll och dess språkliga form.

Undervisningen sker i form av föreläsningar, seminarier, gruppövningar.

Lärandemål

Efter avslutad delkurs förväntas de studerande

- ha en översikt över modern spanskspråkig litteratur från 1900-talet
- kunna identifiera olika typer av texter och litterära genrer i modern spanska
- kunna göra enkla analyser på skönlitterära texter
- kunna redogöra i tal och skrift för ingående skönlitterära texters innehåll och språklig form

Examination

Delkursen examineras genom ett individuellt skriftligt arbete, resuméer av valda skönlitterära texter och aktivt deltagande i seminarierna. För betyget Godkänd på delkursen krävs att alla delar är fullgjorda och godkända. För Väl Godkänd krävs dessutom att det individuella skriftliga arbetet har betyget Väl Godkänd.

Förkunskapskrav

För tillträde till kursen krävs grundläggande behörighet för högskolestudier samt SvB/Sv2B och Spanska lägst C-språk B/steg 3 med lägst betyget Godkänd eller motsvarande kunskaper.

Kursuppläggning/lärandets former

För lärandets former se delkurserna.

Examination och betyg

Kursen bedöms med betygen Väl godkänd, Godkänd och Underkänd. För att få betyget Väl godkänd krävs Väl godkänd på minst tre av delkurserna. För examinationsformer se varje delkurs.

Kursvärdering

Kontinuerlig kursutvärdering under kursen gång. Sammantagen skriftlig utvärdering på hela kursen före dess slut.

Kurslitteratur

Kurslitteratur förtecknas i särskild bilaga.

Övrigt

Undervisningen bedrivs delvis eller helt på målspråket.

Anexo 3. Carta de presentación de la tesis a los participantes.



GÖTEBORGS UNIVERSITET

Institutionen för språk och litteraturer

Avdelningen för spanska

¡Gracias por leer esta información!

Como parte de mis estudios de doctorado, trabajaré en una tesis que se orienta a profundizar principalmente en la lengua escrita de estudiantes universitarios suecos de español como lengua extranjera (E/LE). La metodología comprende el análisis de textos con ayuda de un procesador automatizado. A través de nuestro estudio pretendo identificar mejor algunos aspectos gramaticales que presentan dificultades para estudiantes suecos de E/LE, para de esta forma preparar materiales más efectivos para la progresión en la lengua extranjera.

Te estaría muy agradecida si quisieras colaborar voluntariamente con mi estudio. Tu colaboración consistiría en responder a una pequeña encuesta y enviarme por correo electrónico las tareas de redacción correspondientes al curso *Spanska, Grundkurs I (Del Språkfärdighet/Destrezas)* durante el otoño de 2009.

Siguiendo las normas éticas se garantiza el anonimato de los participantes, así como que la información recibida solamente será utilizada con fines de investigación.

Si tuviesen alguna duda, por favor, pueden contactarme personalmente en la facultad o por correo electrónico: ayme.pino.rodriguez@rom.gu.se

También, si así lo desean, pueden contactar a mis tutores.

Ken Benson: Ken.Benson@rom.gu.se

Isabel Verdaguer Clavera: i.verdaguer@ub.edu

Anexo 4. Certificado de consentimiento informado.

Certificado de consentimiento informado

Yo (escribe tú nombre): _____

He sido informado de la finalidad de la tesis doctoral *Variaciones en contexto: un análisis del uso de las palabras que y aunque en la producción académica escrita de aprendices suecos de E/LE* cuya autora es Aymé Pino, y de su rigurosidad ética, y doy mi consentimiento informado para que los textos por mi elaborados y enviados para este fin sean utilizados en la mencionada tesis.

Fecha: _____

Firma: _____

Anexo 5. La encuesta

ENCUESTA

¡Contesta o marca con una X donde corresponda!

1. Datos del participante

a. Edad: _____

b. Sexo: _____

Femenino

Masculino

c. Lengua materna:

sueco _____

español: _____

otra (especifique): _____

2. ¿Cuánto tiempo llevas estudiando español como lengua extranjera?

a. Menos de 2 años

b. 2-4 años

c. 5-6 años

d. Más de 6 años

3. ¿Dónde has cursado estudios de español como lengua extranjera?

a. Solamente en Suecia _____

b. En Suecia y otro(s) país(es) no hispanohablante(s). Especifica en cuál(es):

c. En España

d. En un país de América Latina

4. Si has cursado estudios de español en el extranjero, especifica por favor, durante cuánto tiempo

a. Menos de 2 años

b. 2-4 años

c. 5-6 años

d. Más de 6 años

4.1 ¿Con qué frecuencia estudiaste español en el extranjero?

a. 20 horas semanales o más

b. Entre 10 y 20 horas semanales

c. Entre 5 y 9 horas semanales

d. Menos de 5 horas semanales

5. ¿Has estudiado otra(s) lengua(s) extranjeras?

a. Sí

b. No

5.1 En caso de respuesta afirmativa a la pregunta anterior, especifica, por favor, la lengua(s)

- Inglés
- Francés
- Alemán
- Italiano
- Portugués
- Otra/s _____

5.2 ¿Cuánto tiempo has estudiado la(s) lengua(s) a que se refiere la pregunta número 5?

- a. Menos de 2 años
- b. 2-4 años
- c. 5-6 años
- d. Más de 6 años

6. Autovaloración de las destrezas escritas en lengua española

a. ¿Cómo valoras tu comprensión de lectura en español como lengua extranjera?

- Excelente
- Muy bien
- Bien

Regular

Mal

b. ¿Cómo valoras tu expresión escrita (redacción) en español como lengua extranjera?

Excelente

Muy bien

Bien

Regular

Mal

¡Gracias!/ Tack!