

**Efectos del traductor de Google sobre la diversidad léxica: el desarrollo de vocabulario entre estudiantes de español como lengua extranjera**

***Effects of Google translate on lexical diversity: vocabulary development among learners of Spanish as a foreign language***

**Kent Fredholm**

*Karlstad University and Uppsala University (Sweden)*  
Kent.Fredholm@kau.se, Kent.Fredholm@moderna.uu.se

**RESUMEN**

Este artículo presenta los resultados de un estudio longitudinal sobre el desarrollo de vocabulario entre un grupo de estudiantes suecos de español como lengua extranjera. Durante un año escolar, los estudiantes (N=31) escribieron cuatro redacciones en español, una prueba previa y una prueba posterior. 15 estudiantes usaron un diccionario impreso como herramienta de traducción, 16 el Traductor de Google (GT). El desarrollo de vocabulario escrito fue medido como el cambio de diversidad léxica entre la prueba previa y la prueba posterior, utilizando el índice de Guiraud. Los resultados muestran que el uso de GT da una diversidad léxica mayor mientras los alumnos lo usan, un efecto que desaparece cuando GT no se utiliza más. Los resultados indican la necesidad de una comprensión más profunda de estructuras lingüísticas y de léxico, un enfoque reforzado en el vocabulario en la enseñanza de idiomas extranjeros, y una extensión más amplia de enseñanza de herramientas de traducción.

Palabras clave: Diversidad léxica, traducción automática en línea, el Traductor de Google, escritura en idiomas extranjeros, desarrollo de vocabulario.

**ABSTRACT**

*This article presents the results of a longitudinal study of vocabulary development among Swedish upper secondary school pupils of Spanish as a foreign language. During one schoolyear, the pupils (N=31) wrote four essays in Spanish, a pre-test and a post-test. 15 pupils used a printed dictionary as translation tool, 16 used Google Translate (GT). Vocabulary development was measured as the change in lexical diversity from pre-test to post-test, using Guiraud's index. The results show that the use of GT leads to higher degrees of lexical diversity as long as it is used, but that the effect vanishes when the tool is no longer used. The results point to the need for a deeper understanding of language structure and lexicon, a reinforced focus on vocabulary in foreign language teaching, and a widened range of explicit instruction of translation tools and strategies.*

*Keywords: Lexical diversity, Free online machine translation, Google Translate, Foreign language writing, Vocabulary development*

## 1. INTRODUCTION

The present article is part of a longitudinal study positioning itself within the fields of computer-assisted language learning (CALL) and foreign language writing, focussing on changes in writing practices and how writing practices affect language learning. One current change in the school context is the transition from writing by hand to using computers, and the different affordances (cf. Gibson, 1986) this way of writing entails for foreign language learners. From primary education to upper secondary education, Swedish schools are undergoing a rapid and extensive process of digitalisation, a topic of much debate. The Swedish national board of education, Skolverket, emphasises the need for further digitalisation as a way to prepare pupils for, among other things, future professional and educational needs, for creative problem solving, and for enhancing equal possibilities for all pupils regardless of parents' socio-economic or academic background (see Skolverket 2018a, for a collection of resources on digitalisation from the Swedish national board of education; cf. also Swedish Government, 2017). A majority of Swedish upper secondary school pupils now have access to a personal school laptop (Swedish Government, 2017).

Studies of how this widespread access to computers affects teaching and learning in Swedish classrooms are still relatively few, especially concerning foreign language education. Studies on longitudinal effects of computer use on language learning outcomes are relatively scarce also internationally (Svensson, 2008; Buckingham, 2011; Cobo Romani & Moravec, 2011; von Schantz Lundgren & Lundgren, 2011; Livingstone, 2012; Fleischer, 2012, 2013; Grönlund, 2014; Grönlund, Andersson, & Wiklund, 2014). New studies of computer use in foreign language classrooms are therefore of great relevance to language education research and to practising language teachers who try to find their way in today's digitalised classroom environment.

A report on research outcomes on school digitalisation made for the Swedish parliament (Riksdagsförvaltningen, 2016) concludes that international research on computer use in schools shows effects on student engagement and motivation, but that effects on learning outcomes are difficult to find, apart from a few studies in Maine where results in mathematics and essay writing were improved among thirteen-year-olds who used a computer (Riksdags-förvaltningen, 2016). Rosén (2012) has also showed that an increase in computer-use among children has led to a decrease in written comprehension, especially among the most prolific computer-users.

Swedish studies (Grönlund, 2014) show that computer use may reinforce achievement patterns already existing in a school and that this might also increase social discrepancies among pupils, rather than reducing them. In a report on the

effectiveness of the use of ICT (information and communication technology) in education, Haelermans (2017) writes that “computer-directed instruction” (p. 34) can improve learning in mathematics and language to a small extent, but that effects vary between students, depending on such things as personal preferences, and she clarifies that results from different studies are not conclusive. Especially regarding language learning, Haelermans writes that studies on ICT effects are few and that very few of these studies have found that ICT use enhances language learning outcomes. Ghysels & Haelermans (2018) find enhanced spelling performance among 7<sup>th</sup> grade pupils in the Netherlands, especially among low-performing pupils. As seems to be the case also with the studies reported in Haelermans (2017), the effects appear to be the result of software enabling frequent repetition of and feedback on easily automatised, basic skills such as spelling. Most studies find effects only on mathematics and only in developing countries, whereas indications of effects on language learning are scarce. In an experiment in the Netherlands, Haelermans (2017) finds effects on spelling and grammar in 7<sup>th</sup> and 8<sup>th</sup> grade. Over all, ICT used as a complement to the teacher seems to yield best results, albeit the effects reported are very small.

In earlier studies of laptop use among Swedish upper secondary school pupils studying Spanish as a foreign language, Fredholm (2015a, 2015b) observed a widespread use of Google Translate (GT) and similar kinds of free online machine translation (FOMT) when the participating pupils used their school laptops to write essays in Spanish. When compared to essays written without FOMT, the use of GT led above all to fewer errors in spelling and article/noun/adjective agreement; at the same time, however, the use of GT increased the number of syntactic errors in the pupils’ essays. The latter effects were the opposite of what the pupils themselves believed would be the outcome.

Fredholm (2015b) focusses on complexity, accuracy and fluency in Spanish texts written with FOMT, but also finds a possible effect on lexical diversity, with slightly higher diversity among FOMT-using pupils, however without investigating whether this effect lasts over time. The present article studies changes in lexical diversity over one schoolyear in closer detail, observing lexical diversity at each writing session and the development from the start to the end of the schoolyear. As in Fredholm (2015b), comparisons are made to pupils using printed dictionaries, but here, an attempt to pinpoint learning outcomes is made by the use of pre- and post-tests. The article contributes, thus, to deepened insights in digitalised foreign language writing practices, focussing on longitudinal effects on vocabulary learning of the regular use of GT as a translation tool in essay writing. The article does so by presenting longitudinal observations of lexical diversity in essays written with GT, comparing with the lexical diversity in essays written with printed dictionaries as a translation tool. A pre- and post-test design enables comparisons from the start to the end of a schoolyear and reveals both differences and similarities between the googling and non-googling participants of the study.

## 1.1 A note on terminology

*Machine translation* is often used as an umbrella term for digital tools for professional translators, as well as for free machine translation services available on the Internet,

such as GT, Babelfish etc. Other terms such as *automatic translation*, *online translation* etc. are also used in the literature. (See O'Neill, 2012, for a discussion on terminology.) When talking about freely available translation sites such as GT, *free online machine translation* (FOMT) probably is the most precise and accurate term. To avoid confusion, the term FOMT will be used throughout the present article, also when referring to studies using any of the above mentioned terms.

## **2. AIM AND RESEARCH QUESTIONS**

The aim of the present study is to investigate the development of upper secondary school pupils' Spanish vocabulary in writing, understood as changes in lexical diversity produced in essays written during a schoolyear, comparing possible differences in lexical diversity between pupils using GT and pupils using printed dictionaries as translation tools. The study observes translation practices and their effects on lexical diversity in texts written by upper secondary school pupils in Sweden studying Spanish as a foreign language on intermediate level (CEFR levels A2.2-B1.1). The following research questions are asked:

- 1) Does the use of GT have any effects on lexical diversity in texts written in Spanish, compared to texts written with printed dictionaries? If so, what are these effects?
- 2) If differences can be found in lexical diversity depending on which translation tool has been used, does the repeated use of GT during one schoolyear have any lasting effects on pupils' development of lexical diversity, as compared to pupils using printed dictionaries? If so, what are these effects?

The present study differs from earlier studies such as Fredholm (2014), (2015a) and (2015b) by focussing exclusively on variation in vocabulary and over a more sustained period of time (one schoolyear rather than three months). Whereas lexical diversity is approached also in Fredholm (2015b), it is examined in greater detail in the present study, offering insights into fluctuations in pupils' vocabulary range over time. Fredholm (2015a) discusses pupils' digital writing strategies, a topic that will not be investigated here.

## **3. THEORETICAL BACKGROUND**

### **3.1 Foreign language writing and lexical diversity**

Researchers such as Harklau (2002), Swain (1996, 2005), and Swain & Lapkin (1995) talk about the importance of productive output for language learning and the role of

writing in a foreign language in order to strengthen the learning of the language (see Manchón, 2011, for an exposé of research on foreign language writing research). Writing enables learners to notice “gaps in their interlanguage” (Manchón, 2011:47; cf. Izumi, 2003). In line with these theories, foreign language writing could be an efficient way for learners to notice instances when their own vocabulary is insufficient, and to give them opportunities to work with and expand their vocabulary and make it richer or more varied. Vocabulary size is important for performance in all language skills (Milton, 2013) and studies of (effects on) vocabulary development are therefore of interest to researchers and practicing foreign language teachers alike.

Research on vocabulary development and word retention in foreign language learning underpins the necessity for repetition of new words in order for them to be retained in long-term memory. Schuetze (2015), Schmitt (2008), and Schuetze & Weimer-Stuckmann (2011) write that suggestions between three and twenty times have been mentioned in earlier studies, Schuetze (2015:38) stating that five times “would be a good way to start” for English speakers learning German. A uniform spacing of word repetition is preferable to a more varied pattern, for retention of vocabulary (Schuetze & Weimer-Stuckmann, 2011). Words need to be encountered several times over extended time periods. It is reasonable to think that regular use of GT or similar translation tools could give language learners opportunities to encounter new words more frequently than what may be the case with printed dictionaries, as the use of GT often is perceived as easier and faster and can give access to more frequent vocabulary encounters (cf. Fredholm, 2015a).

In Fredholm (2015b), however, lexical diversity (used as a measure of lexical complexity) was identical or nearly identical in texts written with FOMT and with printed dictionaries, and individual variations in lexical diversity were correlated to pupils’ grades rather than to the use of translation method. The study reported in Fredholm (2015b) ran over three months. More frequent input over a longer period of time might have stronger effects on vocabulary learning, and the study in Fredholm (2015b) is here repeated with a higher number of texts written during an entire schoolyear, enabling a more detailed picture of changes in written lexical diversity over a longer period of time.

Lexical diversity is a common measure of vocabulary size or richness. According to Milton (2009:127), more able language learners are likely to produce texts with a higher degree of lexical diversity, or, to put it another way, word variation. Malvern & Richards (2002:87) define lexical diversity as “the variety of active vocabulary deployed by a speaker or writer”. Richness and complexity of vocabulary is a good indicator of proficiency in a foreign language (Daller & Xue, 2007) and of great importance for the ability to communicate in another language (Levitzky-Aviad & Laufer, 2013). Indeed, according to Lindqvist (2016), vocabulary is more important for communication than syntax or morphology. Developing a varied and functional vocabulary, thus, is essential for communication in any language, and poses a daunting task to foreign language learners. In their study on French lexical proficiency development over time, Bulté, Housen, Pierrard & Daele (2008) clarify the need for a well-developed vocabulary for language proficiency and underline the dearth of longitudinal research on learning processes and factors impacting on vocabulary learning.

There are few studies of lexical diversity in the Swedish school context, as Berton (2014) points out in his Master's thesis. His study on lexical richness in written production of Spanish as a foreign language shows that language proficiency affects lexical diversity. Berton uses Guiraud's index, among other measures, a method used also in the present article.

Lexical diversity should not be seen as an infallible and all-encompassing measure for proficiency in a foreign language, as a good text draws rather on the good use of vocabulary than on the variety of vocabulary (cf. Malvern et al., 2004, *Introduction*). Different "lexical diversity variables" have, however, been found "to be valid as developmental indices" (Malvern et al., 2004:6. Cf. also Wolfe-Quintero, Inagaki & Kim, 1998:104, talking about lexical variation and sophistication as "related to language development"). A close observation on the development of lexical diversity may thus be seen as an indicator of pupils' language learning, in the present case related to the use or non-use of GT.

### **3.2 Earlier research on the use of free online machine translation in the foreign language writing context**

Up till present day, the number of studies on GT or other kinds of online translation services in foreign language writing research is rather scarce (Thue Vold, 2018), although the number of studies is increasing. Earlier research has mainly studied machine translation use for translator training (Thue Vold, 2018; Gaspari, Almaghout, & Doherty, 2015; Niño, 2009, 2008) and/or FOMT use among university students. As Thue Vold (2018) and Somers (2001) point out, FOMT was not originally created for language learners, but as it is common knowledge that many language learners do use it, it is an important object of investigation.

Thue Vold (2018) offers a valuable review on earlier research on "Machine Translation as a Language Learning Tool" (p. 70), and concludes that "the literature on foreign language teaching and learning provides little evidence that the use of FOMT can be beneficial for language learners in a school context" (p. 72), an assertion also made in Kazemzadeh & Fard Kashani (2014). In earlier research on FOMT in the foreign language learning context, several researchers base their critique, it seems, rather on an intuition that FOMT probably is not good for language learning, rather than empirical findings showing that that is, or is not, the case (Clifford et al., 2013; Somers, 2007; Somers et al., 2006; Steding, 2009; cf. Thue Vold, 2018). Steding (2009) fears that FOMT does not develop learners' own language proficiency, and wants to prevent its use. Somers et al. (2006) write that learners may trust the FOMT output too much, especially at the beginner levels. The findings of O'Neill (2012), studying French learners at university level, point to better intelligibility and more accurate grammar and spelling in texts where FOMT has been used, but no effect on overall text quality, and he concludes that little points to more language learning (or indeed less) among FOMT using students. Likewise, no overall effects of FOMT use pointing conclusively in one direction are seen in Fredholm (2014, 2015a, 2015b), where pupils using FOMT when writing in Spanish produced fewer errors regarding spelling and article/noun/ adjective agreement, but more errors concerning syntax, as compared to pupils using printed dictionaries or no translation help at all, on the contrary to what the participating pupils themselves believed.

On the other hand, researchers such as Jolley & Maimone (2015), O'Neill (2012), and Williams (2006) claim, rather, that FOMT may enhance language learning, provided it be competently used and that language learners have sufficient knowledge about the language that they are learning. Jiménez-Crespo (2018) argues that FOMT has become "another de facto dictionary for language learners" (p. 4), widely used, and that teachers and learners need to understand better how it can and cannot be used. This view is shared by Schnitzer & Gromann (2017) who state that language learners nowadays need to be able to efficiently and competently use the wide array of lexicographic resources available.

A few studies investigating effects of FOMT use show that it may indeed not be all bad; Niño (2008) finds among advanced learners that FOMT can raise language awareness and help with reducing errors. Looking at FOMT use among beginners, Garcia & Pena (2011) find that it helped learners, especially with lower proficiency, to produce longer texts (one might discuss, though, whether this is a good thing per se), texts that received better marks from independent graders. However, judging from pauses and editing interventions, the researchers conclude that the participants probably would learn more from using their own knowledge of the target language, writing without FOMT. Likewise, Kazemzadeh & Fard Kashani (2014), who study Iranian learners of English as a foreign language, find longer and syntactically more complex texts as an effect of FOMT use, especially among beginners, with lower language proficiency.

Thue Vold (2018) studies pupils' metalinguistic talk about the quality of a text automatically translated in French. She finds that FOMT can indeed be used as a means to make pupils discuss metalinguistic issues in foreign language learning. The use of GT as a way to raise metalinguistic awareness is also mentioned by Williams (2006), who states that pupils need to learn how to use FOMT and other tools critically, and by Niño (2008). Another scholar stating that FOMT may enhance some language skills is Giannetti (2016), who has studied writing in Spanish among seventh grade pupils using GT and finds that GT led to fewer errors in syntax and semantics. According to Giannetti, pupils need training in using GT successfully, and insufficient language proficiency made using it more difficult. He argues that GT may "support learning in the foreign language classroom" (p. 4) and that it can build "foreign language literacy" (p. 5), if pupils are instructed how to use it strategically. Finally, Knospe (2017), who studies writing in German as a foreign language among Swedish upper secondary school pupils, concludes, among other findings, in line with Fredholm (2015b), that pupils need a high proficiency in the studied language to be able to make competent use of online resources such as GT, and that it is important that they learn how digital writing tools can support their writing and learning.

## **4. METHODS AND PARTICIPANTS**

### **4.1 Participants**

Two teachers (here renamed as "Carla" and "Belinda") participated with one class each in the study. To simplify the presentation of the data and of the results, the two Spanish classes will be treated as one group in the present article.

The pupils (in total N=31, 26 girls, 5 boys) were randomly divided into two subgroups, here called the "googlers" (N=16, 13 girls, 3 boys) and the "non-googlers" (N=15, 13 girls, 2 boys). A higher number of participants would have been desirable, but practical circumstances such as heavy teacher workloads and unwillingness among pupils to take part in such a demanding, long-term project reduced the number of available participants. During the schoolyear, 18 out of the initially 49 participating pupils chose not to fulfil the study.

The uneven distribution of boys and girls may be considered unfortunate, but reflects the gender distribution at the study programmes where the data collection was performed (the Social science programme and the Humanities programme (Skolverket, 2012)). The random subdivision was controlled for grade levels in Spanish (from the previous schoolyear), study programme, and gender, with some minor alterations done to ensure a more even distribution of, in particular, Spanish grade levels from the previous year among both googlers and non-googlers.

### 4.2 Data collection

The collected data consist of a pre-test consisting of an essay and a grammar test, four intervention essays, a post-test with yet another essay and the same grammar test as in the pre-test, and screen-recordings of the googlers' computer screens. In the present article, the screen-recordings will be used as background information on the amount of text being googled in each intervention essay.

The pre- and post-tests were written by hand, in order to prevent the use of technology playing any part in these results. No dictionaries were allowed and the pupils' mobile phones were collected prior to the pre- and post-tests, as well as before each intervention essay. Each writing session lasted for 50 minutes (a time-frame chosen to give the pupils enough time to write without feeling stressed by too narrow time constraints but also fitting within the pupils' Spanish lessons).

The topics for the four intervention essays (see section 4.3) were elaborated together with Carla and Belinda, the two teachers, in order to fit within the planned syllabus and to relate to text genres mentioned in the national curriculum for foreign languages, level 4 (Skolverket, 2018b). Level 4 in the Swedish national curriculum for foreign languages roughly corresponds to CEFR levels A2.2 to B1.1 (Skolverket, 2017) and was chosen as it is the highest level most upper secondary school pupils reach, and also because the curriculum for this level is particularly rich in text genres, making it suitable for a study on pupils' writing. The all-in-all six writing sessions were evenly distributed during the schoolyear (three each semester, from the pre-test in early September 2016 to the post-test in late May 2017). Carla and Belinda had access to all essays produced during the data collection, and used the texts to give formative feedback to the pupils.

The googlers were permitted Internet access, and were allowed to use GT when needed during the intervention essay writing sessions. The non-googlers were prevented from accessing the Internet, still writing on their laptops but in an online application called *Digiexam*<sup>1</sup>, that blocks Internet access and prevents pupils from fetching previously saved files from the computer. The students were already accustomed to the application, as it is used at their school during written exams in several subjects. The non-googlers were allowed to consult a printed, high quality



Swedish-Spanish/Spanish-Swedish dictionary (Benson, Strandvik, & Santos Melero, 2000), that they had been able to use before during their first year of Spanish studies at the same school.

The pupils were not encouraged to use GT during their regular Spanish lessons with Carla and Belinda (lessons that mainly focused on reading, listening and oral communication). Naturally, every action during the regular Spanish classes cannot be accounted for, and the pupils' use of GT outside the school context cannot be controlled, but the pupils had no Spanish writing tasks assigned as homework, which makes their frequent use of the technology in extramural contexts implausible.

The googlers' computer screens were recorded using the online screencast service of *apowersoft.com*<sup>2</sup>. Two recordings are missing due to technical issues. The screen-recordings are used for detailed analyses of the pupils' interaction with the GT interface (forthcoming), and to register each translation instance made by the pupils, all-in-all 4112 searches ranging from single words over phrases and sentence fragments to complete, complex sentences. The collected data relevant for the present paper are summarised in table I.

	Googlers	Non-googlers	Total
<b>Pre-tests</b>	16	15	31
<b>Intervention essays</b>	59	57	116
<b>Post-tests</b>	16	15	31
<b>Screen recordings</b>	57	-	57
<b>Translation instances</b>	4112	-	4112

Table I. Summary of collected data.

### 4.3 Essay topics

The topics for the pre- and post-test essays and the intervention essays are presented here in abridged form. The instructions were given in Spanish, and each topic was accompanied by a few simple drawings to further inspire the writers. The order in which the intervention essays were written differed between Carla's and Belinda's pupils the second and third writing sessions, as the groups worked with a holiday theme and a fairy-tale theme in different periods during the schoolyear.

#### *Pre-test*

Imagine that you are 85 years old. Think back at your life and write down your memories. You can talk about things like work, good times, bad times, family, travelling, your dreams or things that you regret.

#### *Intervention essay 1 (Belinda's and Carla's pupils)*

Reply to a letter from Pablo, 19, who wants to know what he can do to make his friends take interest in things that he likes, rather than drinking alcohol and hanging about the town.

#### *Intervention essay 2 (Belinda's pupils)/Intervention essay 3 (Carla's pupils)*

Write an argumentative text for or against travelling. Why should we travel, or why should we not travel? You can also write about your favourite destination, and explain why you recommend it to others.

*Intervention essay 3 (Belinda's pupils)/Intervention essay 2 (Carla's pupils)*

Write about the traditions and holidays in your country of origin. You can write, for instance, how different holidays are celebrated, explain the origins of some traditions, tell about traditions that no longer are observed or talk about traditions that you do not like and would like to change.

*Intervention essay 4 (Belinda's and Carla's pupils)*

Imagine that you are Little Red Riding Hood. You are now 75 years old. When your grandchild comes to see you, you retell him the story about your adventure many years ago in the forest, when you met the wolf.

*Post-test*

How was your life when you were a little kid? What were your dreams for the future? What did you usually do? Write about, for instance, your family, your friends, good times, bad times, interests and dreams.

#### 4.4 Analyses and measures

Lexical diversity can be measured in a variety of ways, ranging from the straightforward Type-Token Ratio (TTR) used since 1944 (Johnson, 1944; Daller & Xue, 2007) to more sophisticated measures such as *D* (Daller & Xue, 2007). The type/token ratio is reliable when dealing with texts of equal or very similar lengths; as text length increases, though, the type/token ratio naturally decreases, making the measure unreliable for longer texts and, especially, for texts of highly varying lengths. Several measures based on the TTR have been elaborated to compensate for this, one of the most frequently used being Guiraud's index (Guiraud, 1954; van Hout & Vermeer, 2007; Milton, 2009), also chosen for this study. Guiraud's index uses the formula  $V/\sqrt{N}$  (i.e. Types/ $\sqrt{\text{Tokens}}$ ), dividing the number of types (or lemmas, unique words) in a text, by the square root of the number of tokens (the total number of words in the text). As every measure of lexical diversity, Guiraud's index has been criticised for being sensitive to variations in text length; it is, however, often found to be one of the most reliable measures (especially for shorter texts up to a "few hundred tokens" (Malvern et al., 2004:29), and van Hout & Vermeer (2007) find it often to be the better option between various measures. Guiraud's index was chosen as the most suitable for the present study, as the collected essays are all quite short and of similar lengths. The average length of the essays (number of tokens) and average number of unique words (types) are summarised in table II.

	Googlers		Non-googlers	
	types	tokens	types	tokens
<b>Pre-test</b>	60.75	200.09	57.53	142.00
<b>Intervention essays</b>	86.17	191.59	75.27	139.35
<b>Post-test</b>	97.19	216.00	96.53	220.53

*Table II. Average numbers of types/tokens.*

In the present study, a *word* is defined as a graphic unit of one or more meaning-making letters surrounded by spaces or punctuation marks. Non-Spanish words,

proper names and numbers not written with letters were excluded from the analysis. Inflectional forms of e.g. the same verb or singular and plural forms of the same noun were considered as one lemma or type (cf. Tidball et al., 2007). Bulté et al. (2008) state that only semantic content words should be used when measuring lexical proficiency, as grammatical function words indicate learners' grammatical competence rather than their lexical proficiency. However, as they also point out, most studies mix these word types. The choice was made also in the present study to measure both content words and function words, a choice based on the fact that the participating pupils show a low degree of awareness of grammatical functions of some words; it is plausible that the participating pupils treat all words merely as words, and excluding one kind of words would give an unjust view of their productive written vocabulary.

The development of productive lexical diversity was measured comparing the Guiraud value in the essay part of the pre-test to the Guiraud value in the post-test essay. Statistical significance between the googlers and the non-googlers (with  $p < 0.05$ ) was controlled using a *t*-test in SPSS version 25.

## 5. RESULTS

An overview of the mean level of lexical diversity in the texts from each writing session is presented in section 5.1. These results are further commented in relation to the research questions in section 5.2.

### 5.1 Lexical diversity in the pre-test, post-test and intervention essays.

The mean Guiraud values from each writing session are presented in Table III. Compared to the Guiraud values found in Fredholm (2015b), reaching mean values of 4.08 among pupils using different kinds of FOMT and 3.87 among pupils using printed dictionaries, the values in the present study are higher, which might be explicable by the fact that writing time was shorter (30 minutes) in Fredholm (2015b) than in the present study (50 minutes).

	all		googlers		non-googlers	
	Mean	SD	Mean	SD	Mean	SD
<b>Guiraud pre-test</b>	5.39	0.15	5.46	0.24	5.31	0.20
<b>Guiraud essay 1</b>	5.51	0.13	5.85	0.13	5.17	0.20
<b>Guiraud essay 2</b>	6.19	0.14	6.48	0.19	5.88	0.17
<b>Guiraud essay 3</b>	5.80	0.16	6.25	0.25	5.36	0.15
<b>Guiraud essay 4</b>	5.97	0.17	6.40	0.20	5.51	0.22
<b>Guiraud post-test</b>	6.36	0.27	6.61	0.25	6.47	0.26
<b>Guiraud change from pre-test to post-test</b>	1.16	0.15	1.15	0.19	1.16	0.24

Table III. Guiraud values in pre-tests, intermediate essays and post-tests, and change from pre-test to post-test.

Given the small number of participants, normality testing was done using Shapiro-Wilks for all of the variables in Table III. Normal distribution was found in all cases but one: essay 4 in the googlers' group ( $p$  0.014). An ocular examination of the histogram does however indicate normal distribution or approximate normal distribution. The result is caused by one outlier with a high Guiraud value and two extremes with low Guiraud values; none of these values, however, are exceedingly low or high, compared to Guiraud values in intervention essays 1-3. The result from essay 4 is therefore estimated to be reliable. Equal variance was found, using Levene's test, in all groups at all occasions.

## 5.2 Effects on lexical diversity in the intervention essays and on long-term vocabulary acquisition

The mean Guiraud value from each essay writing session is reported in Figure 1, where the blue line represents the googlers' texts and the orange line the non-googlers'. The values are the same mean values as shown in Table III in section 5.1. The graphs indicate that both groups are making progress in written lexical diversity from the pre-test in September to the post-test in May. Googlers and non-googlers alike show a decrease in lexical diversity in the third intervention essay, at the beginning of the spring term. The reason for this can only be speculated. It is not caused by the essay topic, as Carla's and Belinda's pupils did not write essay topics 2 and 3 in the same order. The temporary decrease in lexical diversity might simply depend on the fact that the pupils had recently been on Christmas holiday for two weeks and had had little contact with the Spanish language during this break, or on other factors beyond the control of this study.

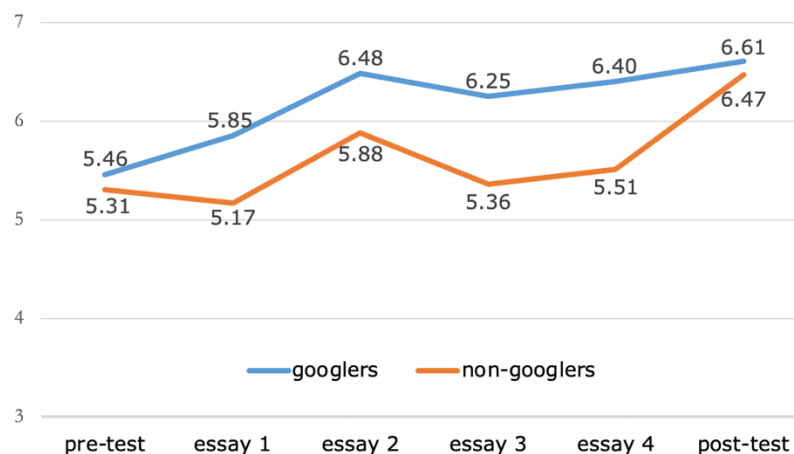


Figure 1. Development of lexical diversity from pre-test to post-test, including the intervention essays.

An independent t-test was conducted to investigate differences between googlers' and non-googlers' mean Guiraud values at each writing occasion, including pre- and post-tests. The results are summarised in Table IV. Levene's test showed that equal variances could be assumed in all cases but one (intervention essay 3,  $p$  0.020). Significant differences were found for every intervention essay, (i.e. for the essays written with GT or dictionaries), but not for the pre-test nor for the post-test. As expected, considering the almost non-existent difference between the groups when comparing the development from the pre-test to the post-test, this difference was also found to be non-significant ( $p$  0.969). The initial hypothesis that the use of GT leads to greater lexical diversity is thus corroborated in as far as the intervention essays; the results do however suggest that this effects is not lasting, but that it disappears when GT is no longer used. This indicates that the use of GT, at least to the extent it was utilised in this study, with high probability does not facilitate the development of a greater productive vocabulary, compared to the use of printed dictionaries (nor, indeed, the other way around).

	t	df	Sig. (2-tailed)
<b>Guiraud value pre-test *</b>	0.475	29	0.638
<b>Guiraud value essay 1 *</b>	2.868	28	0.008
<b>Guiraud value essay 2 *</b>	2.336	27	0.027
<b>Guiraud value essay 3 †</b>	2.336	21.526	0.006
<b>Guiraud value essay 4 *</b>	3.079	27	0.006
<b>Guiraud value post-test *</b>	2.969	29	0.701
<b>Guiraud value change from pre-test to post-test *</b>	0.388	29	0.969

\* Equal variances assumed

† Equal variances not assumed (Levene  $p$  0.020)

Table IV. Differences in mean Guiraud values between googlers and non-googlers.

A crosstab analysis of grade levels and the amount of GT use (Table V) shows that pupils with lower grades generally googled a larger amount of their texts, and that pupils with higher grades googled fewer words (with one exception, a pupil with the grade B who on average googled 64.44% of the words). Again, F, E, and D were counted as low grades, C, B, and A as high. A mean use of GT under the median amount of 37.40% (thus  $<M$ ), was considered as restricted use of GT, whereas a use equal to or above 37.40 % ( $\geq M$ ) was considered as an extensive use of GT. The results show that there is a clear correlation between grade level and amount of GT use (Fisher exact test  $p$  0.001). This reinforces the findings in Fredholm (2015b).

	Restricted use of GT ( $<M$ )	Extensive use of GT ( $\geq M$ )	Total
<b>Low grades (F, E, D)</b>	0	9	9
<b>High grades (C, B, A)</b>	6	1	7
<b>Total</b>	6	10	16

Table V. Crosstab analysis of correlations between GT use and grade level.

When looking at correlations between the amount of googled text in the intervention essays and the Guiraud values in the googlers' essays, no significant correlation can be found (Pearson correlation 0.068). Analyses of correlations between the amount of googled text in the intervention essays and the development of lexical diversity measured as the difference from pre-test to post-test also show non-significant correlation (Pearson -0.056). That is to say, there seems to be no correlation between the amount of GT use and the development of productive lexical diversity. In the present set of data, thus, nothing seems to indicate that more googling would lead to the retention of more words in the long run, albeit claims of causal links cannot be made.

The results reported in this section reflect the statement in Giannetti (2016:18), saying that FOMT "cannot replace proficiency in a language". Using GT may help foreign language pupils to produce texts, but it seems unlikely that it can teach them how to produce texts without it, to follow the line of argument in Garcia & Pena (2011; cf. Giannetti, 2016), who compare GT to a GPS system that enables you to get where you want but does not make you skilled in orientating yourself without help.

## 6. CONCLUSIONS AND DISCUSSION

The present study contributes to research on FOMT use among pre-university level language learners by giving a more detailed insight into the effects, or absence of effects, on vocabulary development of GT use during essay writing in Spanish as a foreign language. In particular, it gives a detailed picture of how lexical diversity may evolve and vary during a schoolyear. As the results indicate that the use of GT does not lead to neither greater nor smaller lasting effects on vocabulary variation, as compared to a more traditional approach, the study may provide foreign language teachers with a more nuanced perception of the role and the affordances of different translation tools.

There are two main conclusions to be drawn from the results in the present study: firstly, that the use of GT does seem to help language learners to write texts with a more varied vocabulary, as long as GT is being used; secondly, that the use of GT does not appear to help the same language learners to gain a more varied active vocabulary over time, when left on their own, as compared to language learners using printed dictionaries. Neither, indeed, does the use of printed dictionaries, in the present study, point to a major increase in productive vocabulary compared to the use of GT. This strengthens the results from Fredholm (2015b), adding a more complex picture of vocabulary development over time.

The same kinds of errors and mistakes were often made by the non-googlers in the present study, albeit they looked up fewer words than their googling peers. Considering observations made in Fredholm (2015a; 2015b), pupils may struggle with understanding results of dictionary searches as well as of GT searches. A higher level of language awareness – both in the mother tongue and in the target language – does seem necessary in order for foreign language learners to be able to use the translation tools at their disposal. Otherwise expressed, in order to become a good foreign

language writer, one needs to reach a high level of knowledge in the studied language, rather than to take shortcuts relying on any kind of translation technology, be it digital or analogue.

As both googlers and non-googlers in the present study gained more or less the same development in written lexical variation at the end of the schoolyear, it may thus be reasonable to conclude that the use of GT or a printed dictionary as translation tool may be of less importance than raising the pupils' awareness of language structure, enhancing their vocabulary as well in the studied language as in their mother tongue, and developing their general linguistic knowledge through explicit instruction. Considering the findings of Levitzky-Aviad & Laufer (2013:144), who show a slow development of active vocabulary in English among Israeli learners, a need for more explicit vocabulary instruction and practice, also of "non-basic vocabulary", in earlier schoolyears, may be necessary.

The notable increase in lexical diversity from the last intervention essay to the post-test, among both googlers and non-googlers, may be puzzling and has no evident explanation. It was the last writing session for the entire schoolyear (not only for the study but for regular lessons as well), and it is plausible that the pupils made an effort to show their best work, as the texts were available also to the grading teachers. Another reason could be that the translation tools (GT and printed dictionaries) used during the intervention essay writing sessions may have been a disturbance to the pupils, leading to a greater focus on translation than on the production of cohesive texts. The absence of the translation tools may, then, have led to a greater focus on the texts themselves. A follow-up study comparing intermediate level learners writing with GT, dictionaries and without any translation tools would be of great interest.

The results of this study are, in a way, double. On the one hand, it is clear that the use of GT does affect lexical diversity, giving the googled texts a wider range of vocabulary. On the other hand, the results also indicate that the effect is immediate but not lasting when GT is no longer used. Considering this, it is reasonable to presume that GT in itself may not have a lasting effect on foreign language vocabulary development. It is also reasonable to conclude that GT is not able to boost less proficient learners' vocabulary; it does help them to write texts, but it does not seem to develop their independency and to become capable of writing lexically more varied texts on their own. Considering what Izumi (2003) writes about output and its ability to make language learners more aware of what they can and cannot express, it seems reasonable to say that this was not the case in the present study, as pupils often did not trust their own knowledge, or their linguistic gut-feeling, if the expression may be allowed.

Foreign language teachers reading this text may wonder what the above-mentioned results might imply for their own teaching. If vocabulary development seems to be the same regardless of pupils' using GT or printed dictionaries, shall pupils be left to choose for themselves? As highlighted in Fredholm (2015a), pupils often struggle with GT and printed dictionaries alike, and need training in both ways of working with translating (cf. Schnitzer & Gromann, 2017, who talk about the need for competence using all available translation resources). Prohibiting pupils from using GT is probably fruitless, and more is gained by showing pupils different ways of using it and discussing together why different translations are more or less trustworthy or

more or less faulty. This may be one way to enhance pupils' metalinguistic reflection, to some extent in line with Thue Vold (2018). This way of dealing with GT in the foreign language classroom setting does not take for granted that its use will enhance pupils' language learning, but views GT, rather, as one of several tools that are available to learners, and that they will benefit from knowing in a less superficial way. GT can be a useful tool to make language learners write, especially less proficient learners who otherwise would not write as much or, perhaps, not at all; it is important, though, that language teachers stress that it is improbable that GT in itself will improve pupils' vocabulary learning.

## NOTES

- 1 <https://digiexam.zendesk.com/hc/sv>
- 2 <https://www.apowersoft.com/free-online-screen-recorder>

## REFERENCES

- Benson, K., Strandvik, I., & Santos Melero, M. E. (eds.). (2000). *Norstedts spanska ordbok*. Stockholm: Norstedts.
- Berton, M. (2014). *La riqueza léxica en la producción escrita de estudiantes suecos de ELE*. Stockholm: Stockholm University. <http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-105624>
- Buckingham, D. (2011). *Beyond Technology: Children's Learning in the Age of Digital Culture*. Cambridge: Polity Press.
- Bulté, B., Housen, A., Pierrard, M., & Daele, S. V. (2008). Investigating lexical proficiency development over time – the case of Dutch-speaking learners of French in Brussels. *Journal of French Language Studies*, 18(3), 277–298. <https://doi.org/10.1017/S0959269508003451>
- Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the Landscape: What is the Role of Machine Translation in Language Learning? *@tic. Revista d'innovació Educativa*. <https://ojs.uv.es/index.php/attic/article/view/2228>
- Cobo Romani, C., & Moravec, J. W. (2011). *Aprendizaje Invisible. Hacia una nueva ecología de la educación*. Barcelona: Col·lecció Transmedia XXI. Laboratori de Mitjans Interactius / Publicacions i Edicions de la Universitat de Barcelona.
- Daller, M. H., & Xue, H. (2007). Lexical richness and the oral proficiency of Chinese EFL students. In H. Daller, J. Milton, & J. Treffers-Daller (eds.), *Modelling and assessing vocabulary knowledge* (150-164). Cambridge, UK: Cambridge University Press.
- Fleischer, H. (2012). What is our current understanding of one-to-one computer projects: A systematic narrative research review. *Educational Research Review*, 7(2), 107-122. <https://doi.org/10.1016/j.edurev.2011.11.004>
- Fleischer, H. (2013). *En elev – en dator: kunskapsbildningens kvalitet och villkor i den datoriserade skolan*. Jönköping University, Jönköping. <http://www.diva-portal.org/smash/record.jsf?pid=diva2:663330>
- Fredholm, K. (2014). Effects of online translation on morphosyntactic and lexical-pragmatic accuracy in essay writing in Spanish as a foreign language (96-101). *CALL Design: Principles and Practice; Proceedings of the 2014 EUROCALL Conference, Groningen, The Netherlands*. Dublin: Research-publishing.net. <https://doi.org/doi:10.14705/rpnet.2014.000201>



- Fredholm, K. (2015a). El uso de traducción automática y otras estrategias de escritura digital en español como lengua extranjera. *Estudios de Lingüística Aplicada*, 33(62), 9-31.
- Fredholm, K. (2015b). Online Translation Use in Spanish as a Foreign Language Essay Writing: Effects on Fluency, Complexity and Accuracy: El uso de traducción automática en la escritura en español como lengua extranjera: Efectos en fluidez, complejidad y corrección. *Revista Nebrija de Lingüística Aplicada*, 2015(18).  
[https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo\\_54ff41bcb4d19.pdf](https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_54ff41bcb4d19.pdf)
- García, I., & Peña, M. I. (2011). Machine translation-assisted language learning: writing for beginners. *Computer Assisted Language Learning*, 24(5), 471-487.  
<http://www.tandfonline.com/doi/abs/10.1080/09588221.2011.582687>
- Gaspari, F., Almaghout, H., & Doherty, S. (2015). A survey of machine translation competences: Insights for translation technology educators and practitioners. *Perspectives*, 23(3), 333-358.
- Ghysels, J., & Haelermans, C. (2018). New evidence on the effect of computerized individualized practice and instruction on language skills. *Journal of Computer Assisted Learning*, 34(4), 440-449.  
<https://doi.org/10.1111/jcal.12248>
- Giannetti, T. R. (2016). *Google Translate as a Resource for Writing: A Study of Error Production in Seventh Grade Spanish* (master of science). School of Arts and Sciences, St. John Fisher College, Pittsford, N.Y.  
[http://fisherpub.sjfc.edu/cgi/viewcontent.cgi?article=1358&context=education\\_ETD\\_master](http://fisherpub.sjfc.edu/cgi/viewcontent.cgi?article=1358&context=education_ETD_master)
- Gibson, J.J. (1986). *The ecological approach to visual perception*. New York, N.Y.: Psychology Press.
- Grönlund, Å. (2014). *Att förändra skolan med teknik: Bortom "en dator per elev"*. Örebro: Örebro university.  
<https://www.diva-portal.org/smash/get/diva2:706366/FULLTEXT01.pdf>
- Grönlund, Å., Andersson, A., & Wiklund, M. (2014). *Unos uno årsrapport 2013*. Örebro university.
- Guiraud, P. (1954). *Les caractères statistiques du vocabulaire. Essai de méthodologie*. Paris: Presses Universitaires de France.
- Haelermans, C. (2017). *Digital Tools in Education. On Usage, Effects and the Role of the Teacher*. Stockholm: SNS Förlag.
- Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, (11), 329-350.
- Izumi, S. (2003). Comprehension and Production Processes in Second Language Learning: In Search of the Psycholinguistic Rationale of the Output Hypothesis. *Applied Linguistics*, 24(2), 168-196.
- Jiménez-Crespo, M. A. (2018). The role of translation technologies in Spanish language learning. *Journal of Spanish Language Teaching*, 0(0), 1-13.  
<https://doi.org/10.1080/23247797.2017.1408949>
- Johnson, W. (1944). Studies in language behavior: I. A program for research. *Psychological Monographs*, 56, 1-15.
- Jolley, J. R., & Maimone, L. (2015). Free online machine translation: use and perceptions by Spanish students and instructors. *Learn Language, Explore Cultures, TransformLives*, 2017.
- Kazemzadeh, A. A., & Fard Kashani, A. (2014). The effect of computer-assisted translation on L2 learners' mastery of writing. *International Journal of Research Studies in Language Learning*, 3(3), 29-44.

- Knospe, Y. (2017). *Writing in a Third Language. A Study of Upper Secondary Students' Texts, Writing Processes and Metacognition*. Universiteit Antwerpen, Umeå University, Umeå.
- Levitzky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight years of study: single words and collocations. In C. Bardel, C. Lindqvist, & B. Laufer (eds.), *L2 vocabulary acquisition, knowledge and use. New perspectives on assessment and corpus analysis* (127-148).
- Lindqvist, C. (2016). Tredjespråkets ordförråd. In C. Bardel, Y. Falk, & C. Lindqvist (eds.), *Tredjespråksinläring*. Lund: Studentlitteratur.
- Livingstone, S. (2012). Critical reflections on the benefits of ICT in education. *Oxford Review of Education*, 38(1), 9-24.
- Malvern, D. D., Richards, B. J., Chipere, N., & Durán, P. (2004). *Lexical Diversity and Language Development: Quantification and Assessment*. Palgrave Macmillan.
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85-104. <https://doi.org/10.1191/0265532202lt221oa>
- Manchón, R. M. (2011). The Language Learning Potential of Writing in Foreign Language Contexts: Lessons from Research. In T. Cimasko & M. Reichelt (eds.), *Foreign Language Writing Instruction: Principles & Practices* (44-64). Anderson, SC: Parlor Press.
- Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Bristol/Buffalo/Toronto: Multilingual Matters.
- Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindqvist, & B. Laufer (eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (57-78). European Second Language Association.
- Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*, 21(1), 29-49. <http://www.tandfonline.com/doi/abs/10.1080/09588220701865482>
- Niño, A. (2009). Machine translation in foreign language learning: language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, 21(2), 241-258. <http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=5579748>
- O'Neill, E. M. (2012). *The Effect of Online Translators on L2 Writing in French*. University of Illinois, Urbana-Champaign.
- Riksdagsförvaltningen. (2016). *Digitaliseringen i skolan - dess påverkan på kvalitet, likvärdighet och resultat i utbildningen Rapport från riksdagen 2015/16:RFR18 - Riksdagen* (No. 2015/16:RFR18). Stockholm: Riksdagsförvaltningen. [https://www.riksdagen.se/sv/dokument-lagar/dokument/rapport-fran-riksdagen/digitaliseringen-i-skolan---dess-paverkan-pa\\_H30WRFR18](https://www.riksdagen.se/sv/dokument-lagar/dokument/rapport-fran-riksdagen/digitaliseringen-i-skolan---dess-paverkan-pa_H30WRFR18)
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language teaching research*, 12(3), 329-363.
- Schnitzer, J., & Gromann, D. (2017). Estrategias de combinación de recursos lexicográficos en la adquisición de lenguas extranjeras (en contextos específicos). *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 11(23), 124-145. <https://doi.org/10.26378/rnlael112334>
- Schuetze, U. (2015). Spacing techniques in second language vocabulary acquisition: Short-term gains vs. long-term memory. *Language Teaching Research*, 19(1), 28-42. <https://doi.org/10.1177/1362168814541726>
- Schuetze, U., & Weimer-Stuckmann, G. (2011). Retention in SLA lexical processing. *CALICO Journal*, 28(2), 460-472.
- Skolverket. (2012). Upper Secondary School 2011. Stockholm: Skolverket. [https://www.skolverket.se/sitevision/proxy/publikationer/svid12\\_5dfee44715d35a5cdfa2899/55935574/wtpub/ws/skolbok/wpubext/trycksak/Blob/pdf2801.pdf?k=2801](https://www.skolverket.se/sitevision/proxy/publikationer/svid12_5dfee44715d35a5cdfa2899/55935574/wtpub/ws/skolbok/wpubext/trycksak/Blob/pdf2801.pdf?k=2801)

- Skolverket. (2017). Kommentarmaterial till kursplanen i moderna språk. Reviderad 2017. Stockholm: Skolverket.  
[https://www.skolverket.se/sitevision/proxy/om-oss/publikationer-och-nyhetsbrev/sok-publikationer/svid12\\_5dfee44715d35a5cdfa2899/55935574/wtpub/ws/skolbok/wpubext/trycksak/Blob/pdf3866.pdf?k=3866](https://www.skolverket.se/sitevision/proxy/om-oss/publikationer-och-nyhetsbrev/sok-publikationer/svid12_5dfee44715d35a5cdfa2899/55935574/wtpub/ws/skolbok/wpubext/trycksak/Blob/pdf3866.pdf?k=3866)
- Skolverket. (2018a). Förändringar och digital kompetens i styrdokument. Stockholm: Skolverket.  
<https://www.skolverket.se/temasidor/digitalisering/digital-kompetens>
- Skolverket. (2018b). Ämne – Moderna språk (Gymnasieskolan). Stockholm: Skolverket.  
<https://www.skolverket.se/undervisning/gymnasieskolan/laroplan-program-och-amnen-i-gymnasieskolan/gymnasieprogrammen/amne?url=1530314731%2Fsyllabuscw%2Fjsp%2Fsubject.htm%3FsubjectCode%3DMOD%26lang%3Dsv%26tos%3Dgy&sv.url=12.5dfee44715d35a5cdfa92a3>
- Somers, H. (2001). Three perspectives on MT in the classroom. I *MT Summit VIII Workshop on Teaching Machine Translation* (25-29). Santiago de Compostela: Citeseer.
- Somers, H. (2007). Machine Translation and the World Wide Web. In K. Ahmad, C. Brewster, & M. Stevenson (eds.), *Words and Intelligence II* (209-233). Springer.
- Somers, H., Gaspari, F., & Niño, A. (2006). Detecting inappropriate use of free online machine translation by language students—A special case of plagiarism detection. I *Proceedings of the Eleventh Annual Conference of the European Association for Machine Translation* (41-48).
- Steding, S. (2009). Machine translation in the German classroom: Detection, reaction, prevention. *Die Unterrichtspraxis/Teaching German*, 42(2), 178-189.  
<http://onlinelibrary.wiley.com/doi/10.1111/j.1756-1221.2009.00052.x/full>
- Svensson, P. (2008). *Språkutbildning i en digital värld Informationsteknik, kommunikation och lärande*. Norstedts Akademiska Förlag.
- Swain, M. (1996). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (eds.), *Principle & Practice in Applied Linguistics* (125-144). Oxford: Oxford University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (ed.), *Handbook of research in second language teaching and learning* (Vol. 1-1, 471-483). Mahwah: Lawrence Erlbaum Associates.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.
- Swedish Government. (2017). Stärkt digital kompetens i skolans styrdokument. Regeringskansliet.  
<https://www.regeringen.se/493c41/contentassets/acd9a3987a8e4619bd6ed95c26ada236/informationsmaterial-starkt-digital-kompetens-i-skolans-styrdokument.pdf>
- Thue Vold, E. (2018). Using Machine-Translated Texts to generate L3 Learners' Metalinguistic Talk. In Å. Haukås, C. Bjørke, & M. Dypedahl (eds.), *Metacognition in Language Learning and Teaching* (67-97). New York: Routledge.
- Tidball, F., Treffers-Daller, J., & Tidball, F. (2007). Exploring measures of vocabulary richness in semi-spontaneous French speech. A quest for the Holy Grail? In J. Treffers-Daller, H. Daller, & J. Milton (eds.), *Modelling and assessing vocabulary knowledge* (133-149). Cambridge, UK: Cambridge University Press.  
<http://www.cambridge.org/catalogue/catalogue.asp?isbn=9780521878517>
- van Hout, R., & Vermeer, A. (2007). Comparing measures of lexical richness. In H. Daller, J. Milton, & J. Treffers-Daller (eds.), *Modelling and Assessing Vocabulary Knowledge* (93-115). Cambridge: Cambridge University Press.

- von Schantz Lundgren, I., & Lundgren, M. (2011). Unga elever med egen dator – Några lärares tankar om hur deras undervisning påverkas. *Utbildning och lärande*, 5(1).
- Williams, L. (2006). Web-Based Machine Translation as a Tool for Promoting Electronic Literacy and Language Awareness. *Foreign Language Annals*, 39(4), 565-578.  
<http://onlinelibrary.wiley.com/doi/10.1111/j.1944-9720.2006.tb02276.x/abstract>
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, & Complexity*. Honolulu, Hawai'i: Second Language Teaching & Curriculum Center, University of Hawai'i at Mānoa.