



Informal acquisition of L2 English vocabulary

Exploring the relationship between online out-of-school exposure and words at different frequency levels

Marja-Leena Niitemaa

Doctoral student, Department of English, School of Languages and Translation Studies, University of Turku

maleni@utu.fi

Introduction

Research has confirmed a positive relationship between participation in online out-of-school activities and language skills in L2 English, such as vocabulary knowledge, reading, listening and writing (e.g., Brevik, 2019; Härmälä et al., 2014; Peters, 2018; Sundqvist, 2019). In Europe, the most active daily Internet users are 16–19-year-olds with a high level of formal education. The rate is particularly high in the Nordic countries, where practically all teenagers have Internet access at home and at school (OECD, 2015). In the Nordic context, out-of-school exposure entails that adolescents may choose their online activities and digital communities freely. The most popular activities are similar across Europe: listening to music with English song lyrics, viewing films and multi-episodic series with or without textual aid, browsing the Internet for fun and reading for information, playing computer games, and social networking (Eurostat, 2015; Statista, 2016). However, the time spent on an activity may differ between age groups, as shown in Peters (2018).

English being one of the most widespread languages on the Internet, it is not surprising that adolescents themselves believe that web-based activities help them acquire English vocabulary (e.g., Brevik, 2019; Voulgari & al., 2014). Some learners tend to find learning through exposure so effective that they may lose motivation towards the formal study of English (Sundqvist & Olin-Scheller, 2013). However, incidental word learning may also take place in school as a by-product of information search for digitized school projects (Cabot, 2018; Kumpulainen & Mikkola, 2016).

Finnish context

In Finland, the curriculum specifically requires the use of digital technology. As in the other Nordic countries, digital literacy is integrated in the curricula (e.g., Finnish National Agency of Education, 2017; Nordic@BETT, 2018) so that digital technology is taught as a general skill needed in all subjects as well as a tool for specific tasks. Schools are also prepared to lend a device for personal use as many teaching materials and course tests are digitized, providing content with pictures, video and audio. Moreover, the digitization of the

high-stakes school-leaving tests (matriculation examination) was completed by spring 2019. The test are taken on a Linux operating system so that the candidates can access only the applications and materials that are installed on the system (<https://www.ylioppilastutkinto.fi/en/matriculation-examination/digital-matriculation-examination>). The students take a test in at least four subjects, one of which must be in an L2. Presently it is English for most students. The language tests include listening and reading comprehension, grammar and vocabulary, and written production. By the school-leaving examination, upper secondary school learners are expected to have reached at least the CEFR level B2 in English (Council of Europe, 2001).

Common European Framework of Reference

Instead of giving recommendations for vocabulary sizes at different proficiency levels, the CEFR descriptors at level B2 (Council of Europe, 2018) entail a broad active reading vocabulary so that learners can scan quickly through long and complex texts to locate relevant details, and understand articles and reports concerned with contemporary problems in which the writers adopt stances or viewpoints. These requirements seem to indicate detailed comprehension skills, which are thought to require recognizing around 98% of the lexis in written texts (Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006; Schmitt, 2008; Schmitt et al., 2011). Researchers commonly agree that learning goals this high cannot be reached without out-of-school language exposure. Nation (2015) and Schmitt (2008) suggest that the bulk of the most frequent 2500–3000 words is acquired at school, and recognition of the 5000 most common word families is likely to be the outcome of both formal and informal learning, while low-frequency vocabulary is mostly acquired through out-of-school activities as rare words require far more encounters than language classes can provide.

Previous research has shown that children and adolescents who use English in online out-of-school activities perform better in tests measuring receptive knowledge of high-frequency words (e.g., Hannibal-Jensen, 2017; Sylvén & Sundqvist, 2012) and mid-frequency words (e.g., Peters, 2018; Sundqvist, 2019). However, little is known of how participation in online out-of-school activities is connected to recognition of low-frequency vocabulary as to date the effects of out-of-school English on infrequent vocabulary have been investigated in relation to productive language use, for example, analyzing essays for polysyllabic words (Sundqvist & Wikström, 2015), academic vocabulary (Olsson & Sylvén, 2015) or infrequent words (Sundqvist, 2019). To fill this research gap, the current study examines Finnish upper secondary school learners' out-of-school online activities in relation to receptive vocabulary knowledge of English at high-, mid- and low-frequency levels, and aims to provide more nuanced information on how the impact of online language encounters varies across different word frequency levels. The frequency levels are defined according to Schmitt and Schmitt (2014) as follows: High-frequency vocabulary comprises words from the first thousand to the third thousand frequency bands; Mid-frequency vocabulary refers to lexis between the fourth thousand and eighth thousand most frequent words; Low-frequency vocabulary begins from the ninth thousand frequency band.

Literature review

Incidental learning refers to learning outcomes without a conscious intent to learn, or learning of one thing when the intention is to learn something else (e.g., Laufer & Hulstijn, 2001). Gass (1999) suggests that a considerable part of second-language vocabulary is

acquired as a by-product of other cognitive exercises involving comprehension. Self-chosen online activities include multiple elements that are traditionally thought to promote incidental vocabulary acquisition, such as comprehensible input (Krashen, 1985), language production (Swain, 1985), noticing (Schmidt, 1990), task-induced involvement (Laufer & Hulstijn, 2001), exposure to and frequent repetition of vocabulary (e.g., Eckerth & Tavakoli, 2012), motivation (Dörnyei & Chan, 2013), as well as self-selected topics and materials with meaningful content (Lee & Pulido, 2017). Incidental vocabulary learning may also depend on the length of previous instruction and the number of instruction hours (Peters et al., 2019), cognateness of words (Lindgren & Muñoz, 2013), learners' proficiency level (Godwin-Jones, 2019), and prior vocabulary knowledge (Peters & Webb, 2018).

Research confirms that online out-of-school contacts with English are beneficial for language acquisition in all age groups from young learners (Hannibal-Jensen, 2017; Lindgren & Muñoz, 2013; Sylvén & Sundqvist, 2012) to teenagers (Brevik, 2019; Peters, 2018; Sundqvist, 2019) and adults (Elgort & Warren, 2014; Rankin et al., 2006; Zheng et al., 2015). Previous results also indicate that self-chosen activities involve active participation and language use in the learner's own niche(s) and interaction with peers sharing the same interests without having to stress over formal assessment, and that learning gains are related to the learner's adopting a role of a language user (Cabot, 2018; Brevik, 2019; Hannibal-Jensen, 2017; Jalkanen & Vaarala, 2013). The following sections introduce recent findings on vocabulary gains from popular online out-of-school activities.

Reading online sites

Reading is traditionally thought to be the best way to increase knowledge of infrequent vocabulary (e.g., Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006, 2016). However, there is an ongoing debate whether voluntary reading can ever provide enough exposure to acquire low-frequency lexis, as it would take 1.5–3 million words to meet rare words repeatedly enough (Cobb, 2007, 2016; McQuillan & Krashen, 2008; McQuillan, 2016; Nation, 2014, 2016). However, McQuillan and Krashen (2008) suggest that a learner would encounter circa 1.5 million words in print when reading 20 minutes per day over two years. Currently, exposure to such a massive number of words may be considered feasible as online encounters often include English (Education First, 2017) and most adolescents spend a great deal of time browsing the Internet and reading self-selected pages related to their other activities (OECD, 2015). Statistics show that over 90% of Finnish teenagers search the Internet for information (Statista, 2015). Online reading can also provide repeated exposure to vocabulary as adolescents tend to visit the same websites (Sockett & Toffoli, 2012) and, moreover, vocabulary learning can take place even if the reader is interested in a limited number of genres (Nation 2014).

Recent studies have been conducted mostly among adults reading printed texts (Elgort & Warren, 2014; Pellicer-Sánchez & Schmitt, 2010; González Fernández & Schmitt, 2015), while investigations into online reading among teenagers are scarcer. However, Sundqvist (2009) reports positive correlations between surfing on the Internet and English vocabulary test scores, as teenagers tend to read English texts online and printed text in L1. Similarly, Peters (2018) found that teenagers preferred online reading to printed text. In relation to vocabulary acquisition, she noticed that visiting English websites explained 17% of the variance of the vocabulary test scores. The effect was greater compared to viewing TV and films without subtitles.

Computer games

Playing computer games is not as widespread as other online activities (OECD, 2015). Less than half of adolescents play one-player games and about one-third participate in collaborative online games. This may be dependent on gaming being a highly time-consuming activity, as gamers can easily spend more than twenty hours a week on computer games (Coxhead & Bytheway, 2014; Lan et al., 2015; Zheng et al., 2015).

Performing well in a game requires mastering the key vocabulary, which motivates gamers to take in the essential lexis (Brevik, 2016). To this end, they often need to interact with the gaming community in speaking and/or writing. Although corpora of games vocabulary are not available for analysis in the same way as transcripts for films (Coxhead & Bytheway, 2014), researchers support the assumption that multiplayer games (MPGs), Massively Multiplayer Online games (MMOs) and Massively Multiplayer Online Role-Playing Games (MMORPGs) provide a large input of lexis. Active gamers seem to know English words that other students do not (Coxhead & Bytheway, 2014), or words that their teachers do not recognize (Brevik, 2016), and they also excel at infrequent words in productive vocabulary tests (Sundqvist, 2019). These observations indicate a better command of rare words acquired through frequent long-term gaming. Gaming also includes bodily action both from the players and the avatars they create, which has been found to activate brain resources that are engaged in lexical processing and language comprehension (Lan et al., 2015).

The time spent on gaming has been found to contribute to higher learning outcomes, such as better reading skills in English compared to L1 (Brevik, 2019), scoring high points in vocabulary tests (Sundqvist & Wikström, 2015; Sundqvist, 2019) and acquiring new words during a game (Zheng et al., 2015). In Peters (2018), however, only a weak connection was detected between adolescents' receptive vocabulary knowledge and gaming. This might be explained by the test design in which the types of games were not differentiated.

Audiovisual exposure

Lexical input provided by audiovisual exposure and the role of subtitles (text in L1) and captions (text in L2) in word learning have been examined widely. Webb & Rodgers (2009a, 2009b) have shown that knowledge of the first 3000 most frequent word families enables understanding of English in multi-episodic TV series, and that frequent repetition of vocabulary facilitates learning of new words. TV series are also rich in phraseology. Examining formulaic language on internet television, Lin (2014) showed that drama, comedy and factual genres are close to everyday speech when compared against the spoken BNC corpus. Moreover, hearing the intonation and flow of speech is likely to promote acquisition of formulaic sequences (Lin, 2014; Lin & Siyanova-Chanturia, 2014). Unlike popular TV series, full-length feature films, particularly action, science fiction and war genres, may contain rare words even from the 7th to the 9th thousand frequency levels, but simultaneous exposure to language, gestures, sound and visual clues are likely to facilitate learning and understanding content with rare words (Sweller, 2005).

Subtitles and captions have been found to be beneficial for learning both vocabulary and formulaic sequences (e.g., Peters et al., 2016; Peters & Webb, 2018). Subtitles give the viewers a chance to compare whether what they heard was the same as what was said, and captions combine the spoken words with their written forms. Research among adolescents (Peters et al., 2016) reveals that captions facilitate the recalling of word forms, while subtitles are related to meaning recall. Both learning outcomes are related to the learners' prior vocabulary size. However, acquisition is also possible without subtitles or captions. As

shown in Peters & Webb (2018), watching a full-length documentary without any textual aid may also increase both meaning recall and meaning recognition. In the same vein, Peters (2018) has found positive significant correlations between vocabulary size and viewing TV or films without subtitles.

Music and social networking

Most adolescents download music on the Internet and listen to English songs on a regular basis (e.g., Statista, 2016). As song lyrics tend to be lexically simple and do not include common formulaic language (Lin, 2014), listening to English music has been found to facilitate vocabulary learning among young learners (Kuppens, 2010; Lindgren & Muñoz, 2013) but not among more advanced learners (González Fernández & Schmitt, 2015; Peters, 2018). Moreover, music is often played in the background when being engaged in something else. However, active listening, such as singing along, focusing on pronunciation or word meanings, learning the lyrics by heart or translating them into L1 is likely to promote acquisition (Kerekes, 2014; Sockett & Toffoli, 2012).

Adolescents are also active in online social networks (OECD, 2015). For many, this may even be the most frequent context for reading (Toffoli & Sockett, 2010). However, the findings on language gains are not encouraging. A large-scale study among Spanish 15-year-olds found a significant but weak positive correlation between reading comprehension scores and participating in online social activities (Gill-Flores et al., 2012). In a Finnish study among the same age group (Härmälä et al., 2014), the correlation between social networking and language skills was non-significant. One explanation might be that most teenagers prefer L1 when participating in online social networks. On the other hand, a significant positive connection has been detected between using English collocations and social networking among adults (González Fernández & Schmitt, 2015). Moreover, Vandergriff (2016) points out that social networking may promote awareness of cultural differences and intercultural communication. These skills, however, cannot be measured by traditional vocabulary tests and need further research in relation to online learning.

All things considered, the connection between digital activities and vocabulary learning is a complex issue to examine. Firstly, researchers mostly rely on learners' self-reports. One notable exception is Verspoor et al. (2011) who had a rare opportunity to examine Dutch teenagers in a naturalistic test setting and compare the effect of media exposure between a group with typical media exposure and another group deprived of contact with popular media for religious reasons. The results indicated significantly weaker performance in writing and vocabulary tests among the latter group. Secondly, differentiating the effects of various types of input is difficult. Participation in any activity may combine reading, communicating and audiovisual input in varying proportions, and moreover, adolescents usually have several online interests. As indicated in Peters (2018) and Brevik (2019), learning outcomes are likely to be the sum of all the types of activities the adolescent is engaged in.

Method

The current study examines Finnish upper secondary school learners' out-of-school activities in relation to receptive vocabulary knowledge of English across high-, mid- and low-frequency levels. Low-frequency vocabulary is of special interest as to date the effects of out-of-school English on infrequent vocabulary have been investigated in terms of productive language use (e.g., Sundqvist & Wikström, 2015; Sundqvist, 2019).

Three research questions are addressed:

1. To what extent are Finnish upper secondary school students engaged in online out-of-school activities in English?
2. What are the participants' self-reported experiences of online out-of-school vocabulary acquisition?
3. What is the connection between online out-of-school activities and recognition of English words at high-, mid- and low-frequency levels?

The first two questions approach the topic from a qualitative perspective, while findings on the third question are based on quantitative analyses.

Participants

The research was conducted among 46 students, aged 16–17, at a typical municipally maintained upper secondary school. The participants volunteered to be examined for various English language skills for two academic years. As the study was carried out during classroom hours, a bigger sample was not possible. The participants had come from different comprehensive schools, but their academic background was homogeneous as the intake to the upper secondary school was based on grade point average. At the start of the experiment, the students had taken circa 600 lessons (45 mins.) in L2 English over seven years in comprehensive school and then circa 30 lessons (75 mins.) in the first term at the upper secondary school level.

Questionnaires

To survey the type and the frequency of online out-of-school activities, and the extent of English exposure, the participants filled out a similar online questionnaire at the end of the first term in the new school, and then at the beginning of the spring term in the second year. Finding consistency in the self-reported data over time was thought to enhance credibility and transferability of the results (Mackey & Gass, 2005; Peer et al., 2012).

Before answering the questions, the participants were instructed to think of a typical school week in their lives. They were shown a list of activities and asked to indicate how frequently they were engaged in these activities online and estimate the proportion of English in them. The list included online reading (*Sites*), playing computer games (*Games*), watching films, video clips, multi-episodic series and streaming services (*Films*), listening to music with English song lyrics (*Music*) and communicating on social network sites (*Social Networks*). With respect to *Games*, the participants were also asked to indicate the age they had started gaming and name the games they were used to playing. The frequency of exposure was measured on a three-point Likert scale, the options being *Daily*, *Weekly* or *Rarely*. The proportion of English content was indicated by ticking one of the following options: *All English*, *Half English* or *Less English*. If the participants had activities that were not included in the list, they had a chance to share additional information by answering an open-ended question. Based on the responses, three exposure variables were formed: 1) the type of activity, 2) the frequency of the activity and 3) the extent of English content per activity.

A third questionnaire was sent out in the spring term of the second year to survey the participants' personal experiences of online out-of-school word learning. The participants were asked to evaluate their vocabulary acquisition by choosing one of the following options: *A lot of words*, *Certainly some words*, *Perhaps some words* or *I don't know*. As an

optional open-answer question, the participants were asked to give examples of words they remembered having learned and/or name semantic contexts pertaining to the words they had learned.

Vocabulary recognition tests and data analysis

Receptive vocabulary knowledge was examined using the *Vocabulary Levels Test* (henceforth the *VLT*; Schmitt, Schmitt & Clapham, 2001; Nation, 1983). In the *VLT*, the scores of the 2nd and 3rd thousand frequency bands (2K and 3K) measured recognition of high-frequency vocabulary; the score of the 5th thousand frequency band (5K) was used to measure recognition of mid-frequency vocabulary and the score of the 10th thousand frequency band (10K) was a measure of recognition of low-frequency vocabulary. Each test section included 30 items. The *VLT* was conducted twice using different test version so that the administration was timed to coincide with the two questionnaires about online out-of-school activities.

For statistical analysis, the qualitative information (type and frequency of the activity, and extent of English content) was coded numerically for each participant. The participants' vocabulary recognition scores were correlated to the coded categories (Pearson's r). Simple linear regression was computed to examine effect sizes of the exposure categories for high-, mid- and low-frequency vocabulary. Finally, stepwise regression was run to examine whether different combinations of two activities could demonstrate a larger effect size for the vocabulary scores than a single activity alone. The analyses were computed using SPSS 23.

Results

To what extent are Finnish upper secondary school students engaged in online out-of-school activities in English? (RQ 1)

Surveys one and two showed that the adolescents used the Internet for the same purposes as other European teenagers reporting contact with English through *Music*, *Films*, *Sites*, *Games* and *Social Network Sites*. When answering the open-answer question, no additional activities were reported. Instead, some participants explained that they had less time for activities particularly in the second year, or named multi-episodic series that they followed on a regular basis. In the following, capital letters and italics are used to denote the exposure categories based on the type and the frequency of the activity and the extent of English content per activity.

The first and the second surveys indicated that participation in the activities was consistent over fifteen months (Table 1). In both surveys, more than half of the participants were in contact with *All-English Daily* exposure through *Music* and *Films*, and slightly less than half through *Sites* and *Games*. Overall, the changes were small in the *All-English Daily* categories. The number of participants reading *Sites* increased by 4.4 percentage points in the second year, while exposure to both *Music* and *Games* decreased by 2.2 percentage points, and *Films* by 4.4 percentage points. *Social Networks* provided the smallest proportion of *All-English Daily* content. The only major change was observed in the category of *Social Networks* with *Less Daily English* showing a decrease of around 20 percentage points in the second year.

Table 1. Participation in online out-of-school activities in English in the first and the second year.

English content	Music		Films		Sites		Games		Social Networks	
	1 st year	2 nd year	1 st year	2 nd year	1 st year	2 nd year	1 st year	2 nd year	1 st year	2 nd year
All-English Daily	26 56.5%	25 54.3%	24 52.2%	22 47.8%	19 41.3%	21 45.7%	21 45.7%	20 43.5%	7 15.2%	7 15.2%
Less English Daily	12 26.1%	11 23.9%	6 13%	6 13%	9 19.6%	10 21.7%	0	0	24 52.2%	15 32.6%
All-English Weekly	0	3 6.5%	5 10.9%	5 10.9%	1 2.2%	2 4.3%	4 8.7%	5 10.9%	1 2.2%	1 2.2%
Less English Weekly	4 8.7%	1 2.2%	6 13%	6 13%	8 17.4%	6 13%	1 2.2%	4 8.7%	4 8.7%	11 23.9%
English Rarely	4 8.7%	6 13%	5 10.9%	7 15.2%	9 19.6%	7 15.2%	20 43.5%	17 36.9%	10 21.7%	12 26.1%

Games differed from the other online activities in two respects (Table 1). Firstly, there was no such category as *Games* with *Less English Daily*. Playing computer *Games* included either *Daily* or *Weekly* encounter with *All-English* content or was played with *Less English* content *Rarely*. The gamers reporting *All-English Weekly* gaming explained in the survey that they were in the habit of playing games regularly with *All-English* content *Daily* but had less time for the activity in the second year of their studies. Thus, the category *Games* refers to *All-English Daily/Weekly* exposure. Secondly, eleven gamers reported playing computer games that required extensive English vocabulary¹ and interaction in English with the gaming community, and named, for example, such games as *Civilization V*, *World of Warcraft* and *Runescape*. They had begun regular gaming by the age of 8–10 and, assisted by their family members, three of them had started even earlier. Based on this information, it seemed reasonable to form two separate gaming categories to be used in the statistical analyses in section 4.3. *Games* includes all the participants who reported *All-English Daily/Weekly* gaming without accounting for the types of game, while the category *Games+* refers to *All-English Daily* gamers with long-term frequent experience of lexically demanding games.

What are the participants’ self-reported experiences of online out-of-school vocabulary acquisition? (RQ 2)

The third questionnaire surveyed the participants’ individual experiences of word acquisition through online out-of-school activities. Out of 46 participants, 30 (65%) ticked the option that they had learned *a lot of words*, 14 (31%) participants chose *certainly some words* and two (4%) participants answered that they *may have learned* some words. The latter were students who had reported a small amount of exposure to English. It is noteworthy that no one chose the option *don’t know*.

The subjects were also asked to report what they had learned. The question was marked as optional, as responding to this question was considered difficult to verbalize. Nevertheless, 36 (78%) participants were able to name altogether over 50 specific semantic contexts, which were then grouped under more general topics in Table 2.

1. The types of games included Massively Multiplayer Online Games (MMO), Massively Multiplayer Online Role Playing Games (MMORPG) and Turn Based Strategy Games (TBS).

Table 2. Semantic contexts of words learned through activities.

Context	Number of responses	Context	Number of responses
Current affairs and society	8	Other hobbies	4
Battle and warfare	6	Vehicles and technology	3
Information technology	6	Business and economy	3
Crime and law enforcement	4	Law and legal proceedings	3
Gaming	4	Culture	3
Music	4	History	2

Ten students did not specify a particular semantic area but remembered having learned “a bit of everything”, “slang”, “American words”, or “useful phrases” or “language needed in everyday interaction”. Moreover, eight participants gave examples of newly learned high- and mid-frequency words such as *corner kick*, *defend*, *promptly*, *queue*, *splendid* and *strike*. Low-frequency words were not mentioned.

What is the connection between online out-of-school activities and recognition of English words at high-, mid- and low-frequency levels? (RQ 3)

To answer RQ 3, the exposure categories introduced in section 4.1 were correlated to the vocabulary test scores (the *VLT*) at high-, mid-, and low-frequency levels (Table 3). Regression analyses were then conducted to examine the potential of types of activities to predict the *VLT* scores at different frequency levels (Table 5).

Vocabulary Levels Test scores

Comparing the *VLT* results over time showed that no significant differences were found in the mean scores of high-frequency words (2K–3K), while the scores of mid-frequency (5K) and low-frequency vocabulary (10K) as well as the total score were significantly higher in the second year compared to the first measurement (Table 3).

Table 3. Results of the Vocabulary Levels Test (the *VLT*).

	K2–K3 Max. 60		K5 Max. 30		K10 Max. 30		Total score Max. 120	
	1 st year	2 nd year	1 st year	2 nd year	1 st year	2 nd year	1 st year	2 nd year
Mean (SD)	44.4 (15.5)	45.3 (13.1)	17.2 (8.73)	20.3 (8.2)	10.2 (6.3)	12.7 (8.1)	73.1 (25.7)	77.5 (30.1)
Percent of the maximum	74%	75.5%	57.3%	67.7%	34%	42.3%	60.9%	64.6%

The correlations between *Daily* activities and the *VLT* scores at high-, mid- and low-frequency levels can be examined in Table 4. The entire list also providing the correlations between less exposure to English and the *VLT* is found in the Appendix.

Type of activity

The type of activity affected the connection between vocabulary recognition and exposure to English. In both measurements, *Daily* language encounters through *Films*, *Games* and *Sites* correlated positively with the *VLT* scores. In the first year, *Games* demonstrated the strongest

correlations between vocabulary scores at all the frequency levels. In the second year, more variation was observed: *Sites* showed the strongest correlation between *high-frequency* words, *Games* between *mid-frequency* words and *Games+* between *low-frequency* words. In comparison with the other forms of exposure, the correlation between *Games+* and vocabulary recognition scores was the strongest at *low-frequency* level and the weakest at *high-frequency* level. In contrast, participation in *Social Networks* demonstrated either positive non-significant correlations or negative significant correlations between the *VLT* scores, even though the exposure was *Daily*. With respect to *Music*, the correlations between the *VLT* were positively or negatively non-significant, excluding one significant positive correlation detected between exposure to *Daily All-English Music* and recognition of *Low-frequency* vocabulary in the first year. However, the phenomenon was not repeated in the second year.

Table 4. Correlations between the *VLT* scores and daily contact with English through online out-of-school activities.

Type of activity	Daily English content	High-frequency		Mid-frequency		Low-frequency
		1 st year	2 nd year	1 st year	2 nd year	1 st year
Music	All-English	non-sig.	non-sig.	non-sig.	non-sig.	.323*
	Less English	non-sig.	non-sig.	non-sig.	non-sig.	non-sig.
Films	All-English	.518**	.371*	.485**	.451**	.476**
	Less English	non-sig.	neg. sig.	non-sig.	neg. sig.	non-sig.
Sites	All-English	.548**	.536**	.497**	.524**	.552**
	Less English	.440*	.327*	.327*	.339*	non-sig.
Games	All-English Daily/Weekly	.691**	.514**	.667**	.571**	.560**
Games+	All-English Daily	.520**	.355**	.554**	.468**	.558**
Social Networks	All-English	non-sig.	non-sig.	non-sig.	non-sig.	non-sig.
	Less English	non-sig.	non-sig.	non-sig.	non-sig.	non-sig.

* $p \leq .050$; ** $p \leq .010$

Frequency (Daily, Weekly, Rarely)

At all frequency levels, significant positive correlations were detected between vocabulary recognition scores and exposure through *Films*, *Sites* and *Games* on condition that the encounters were *Daily*. If the participants were engaged in the activities *Weekly* or *Rarely*, the correlations between the activities and the *VLT* were non-significant or negatively significant (see the Appendix).

English content

The proportion of English content and the frequency of the activity were interconnected. The analyses showed that the correlations between *Films* and the *VLT*, and *Sites* and the *VLT* were positive and significant if the participants were engaged with *All-English Daily* content. If the language contact was *Daily*, but included *Less English*, the correlations were positively or negatively non-significant (*Sites*), or negatively significant or non-significant (*Films*). *Games* was the only activity in which combining *Daily* and *Weekly* exposure with *All-English* content raised the correlation coefficient.

Regression analyses

The results of Simple regression are found in Table 5. In the first year, the largest effect on *high-frequency vocabulary* was found for *Games* (48%), while *Films* (27%), *Games+* (27%) and *Sites* (30%) exhibited an almost equal impact. With respect to *mid-frequency vocabulary*, the largest effects were found for *Games* (45%) and *Games+* (31%), while smaller effect sizes were observed for *Films* (24%) and *Sites* (25%). Equally large effects (31%) on *low-frequency vocabulary* were found for *Games*, *Games+* and *Sites*, while *Films* explained 23% of the variance of the scores for infrequent words.

Table 5. Simple Regression to predict the VLT scores across frequency levels.

High-frequency vocabulary		<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	R2	Adjusted R2
1st year	Films	13.390	3.332	.518	4.019	.000	.269	.252
	Sites	14.366	3.306	.548	4.346	.000	.300	.284
	Games	17.909	2.823	.691	6.343	.000	.478	.466
	Games+	15.730	3.897	.520	4.036	.000	.270	.254
2nd year	Films	11.360	4.287	.371	2.650	.011	.138	.118
	Sites	16.453	3.909	.536	4.209	.000	.287	.271
	Games	15.790	3.971	.514	3.976	.000	.264	.248
	Games+	12.748	5.05	.355	2.523	.015	.174	.155
Mid-frequency vocabulary		<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	R2	Adjusted R2
1st year	Films	8.390	2.280	.485	3.680	.001	.235	.218
	Sites	8.723	2.295	.497	3.801	.000	.247	.230
	Games	11.575	1.947	.667	5.945	.000	.445	.433
	Games+	11.213	2.543	.554	4.410	.000	.307	.291
2nd year	Films	7.345	2.192	.451	3.351	.002	.203	.185
	Sites	8.552	2.098	.524	4.076	.000	.274	.258
	Games	9.322	2.022	.571	4.609	.000	.326	.310
	Games+	8.919	2.542	.468	3.509	.001	.219	.201
Low-frequency vocabulary		<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	R2	Adjusted R2
1st year	Films	5.955	1.657	.476	3.593	.001	.227	.209
	Sites	7.004	1.594	.552	4.394	.000	.305	.289
	Games	7.027	1.565	.560	4.489	.000	.314	.299
	Games+	8.164	1.832	.558	4.456	.000	.311	.295
2nd year	Films	5.633	2.266	.351	2.485	.017	.123	.103
	Sites	6.773	2.202	.421	3.076	.004	.177	.158
	Games	6.282	2.235	.390	4.609	.007	.152	.133
	Games+	11.210	2.275	.596	4.928	.000	.356	.341

In the second year, the largest effect on *high-frequency vocabulary* was found for *Sites* (29%), followed by *Games* (26%), while the effects of *Films* (14%) and *Games+* (17%) remained smaller. Regarding *mid-frequency vocabulary*, the largest effects were found for *Games* (33%) and *Sites* (27%) while *Films* (20%) and *Games+* (22%) showed almost an equal effect. The largest effect on *low-frequency vocabulary* was observed for *Games+*, explaining 36% of the variance of the scores for rare words. Smaller effects were found for *Sites* (18%), *Films* (12%) and *Games* (15%).

In sum, online out-of-school language contacts through *Films*, *Sites*, *Games* and *Games+* promoted learning English vocabulary across all the frequency levels, but the effects of the types of activities varied over time. With respect to low-frequency vocabulary, *Games*, *Games+* and *Sites* were equally good predictors of the score of infrequent words in the first year, while *Games+* had the largest effect on low-frequency scores in the second year.

Table 6. Summary of Stepwise regression analysis for *Sites* and *Games+*.

Model	Unstandardized Coefficients		Standardized Coefficients			
	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	R2
1 (Constant)	14.239	1.488		6.426	.000	
<i>Sites Daily all-English</i>	6.754	2.202	.421	3.076	.004	
2 (Constant)	7.957	1.279		6.221	.000	
<i>Sites Daily all-English</i>	5.037	1.860	.313	2.707	.010	.177
<i>Games+</i>	10.020	2.172	.533	4.612	.000	.449

Finally, stepwise regression was run to examine the combined effect of two activities on *low-frequency vocabulary* (Table 6). The analyses yielded one significant model showing that online reading (*Sites*) and playing lexically demanding computer games (*Games+*) together explained 45% of the variance of the score of *low-frequency* words in the second year. The analyses with other combinations of activities did not meet all the necessary assumptions of regression.

Discussion and conclusion

The present study examined Finnish upper secondary school students' participation in online out-of-school activities in relation to receptive language skills in English. Unlike in previous research (e.g., Peters, 2018; Sundqvist, 2019), the relationship between activities and vocabulary knowledge was analyzed across high-, mid- and low-frequency word levels.

The first research question asked to what extent Finnish upper secondary school students encountered English through online out-of-school activities. The findings showed that roughly half of the students were in daily contact with all-English content through *Music*, *Films*, *Sites* and *Games*, while only a minority used English when participating in *Social Networks*. Furthermore, the changes in participation in the activities surveyed were small over time. Interestingly, reading online *Sites* became slightly more popular than viewing *Films* in the second year. The finding is in line with statistics (Statista, 2015) showing that Finnish adolescents are active seekers of online information.

The second research question inquired into the participants' opinions on acquiring English from online out-of-school activities. Regardless of lexical skills measured by the *VLT*, the participants believed that they had learned English words and phrases through self-chosen online activities and recalled examples of newly learned words. Overall, the findings align with recent research in that the participants considered themselves active users of English within their own digital niches (Cabot, 2018; Brevik, 2019; Sundqvist, 2019). For example, a football enthusiast scoring from low to medium points in the *VLT* followed the careers of his favorite players on English websites, played a football-oriented computer game and reported learning words related to football.

The third research question asked how the type and frequency of the activity, and the extent of English content were related to the vocabulary recognition scores across frequency levels (see Table 4). Regarding *Music* and *Social Networks*, the type of activity mattered most as the correlations between the *VLT* scores were non-significant even if the language contact was daily and the content was in English. However, one positive correlation was found between *Daily All-English Music* and recognition of low-frequency vocabulary in the first year. As types of music were not differentiated, it can only be speculated that some participants may have been interested in genres favoring rare words in the lyrics.

Regarding *Films* and *Sites*, the frequency and the amount of English content were decisive. The correlations between these activities and the vocabulary scores were positive and significant on condition that the participants were in daily contact with all-English content. Moreover, some participants remembered learning useful phrases, which may refer to acquisition from *Films* (cf., Lin, 2014; Lin & Siyanova-Chanturia, 2014). Phraseological knowledge, however, was not tested in the present study.

With respect to *Games*, all three exposure variables were at play. As a type of activity, *Games* required quick comprehension skills, extensive vocabulary and language use in communication (cf., Brevik, 2019). Moreover, as shown in the self-reports, gaming was typically a frequent long-term activity providing a large amount of English. Particularly the participants in the *Games+* category had several years of experience of lexically challenging games. Consequently, the correlations between vocabulary scores and gaming tended to be stronger compared to the correlations between other activities.

As a new contribution to research, the present study showed that the connection between activities and vocabulary varied over time across word frequency levels (Table 5). In the first year, *Games* explained 48% of the variance of vocabulary scores at high-frequency level and 45% at mid-frequency level, while *Sites*, *Games* and *Games+* had an equally large effect (31%) on low-frequency scores. In the second year, the results showed more variation: *Sites* was the best predictor of the variance of high-frequency scores (29%), *Games* had the largest effect on mid-frequency scores (33%). Interestingly, *Games+* predicted 36% of the variance of the low-frequency scores, while the percentage for *Games* was only 15%, which was on a par with *Films* and *Sites*. This indicates that the participants in the *Games+* category were able to improve their knowledge of infrequent vocabulary thanks to long-term frequent language use through gaming, which aligns with findings that vocabulary gains depend on the learner's prior lexical knowledge (e.g., Godwin-Jones, 2019; Peters & Webb, 2018). In comparison with recent research, Sundqvist (2019) found that the type of game appeared to be a less important factor in vocabulary acquisition than the time played, while in the present study, vocabulary gains were connected both with the type of game and the frequency and years of gaming. According to Sundqvist, this issue needs further research. Moreover, the present results showed that reading is an effective way to promote learning of rare words in an online mode, too. Together with playing lexically

demanding computer games, reading online sites explained 45% of the variance of the score of low-frequency words in the second year.

Limitations and suggestions for future research

There are some limitations to the present study. Firstly, the study would have benefited from conducting interviews with the students after the surveys. Our study, however, did not have enough classroom time for that. Secondly, a larger number of participants would have enabled to elicit more nuanced information, e.g., on vocabulary gains through different viewing practices, such as watching one episode of a series at a time or devouring the whole set of episodes at once. Popular films and series may also induce viewers to read the book that the film was based on. For example, an informal conversation with a group of *Hunger Games* fans revealed that they had not only seen the films and followed the English website providing videos, games and information, but also read the printed books in English before translations were available in L1. Similarly, story-driven games may arouse gamers' interest in reading printed books, as shown in Brevik (2019). Investigations among internet subgroups are also needed to uncover how different combinations of online input affect vocabulary acquisition. Future research could also examine learners as producers of English content when participating in different online activities.

Pedagogical implications

Pedagogical implications of the present study relate to the interplay between students' experiences in informal learning environments and motivation to formal learning (Brevik, 2016, 2019; Sundqvist & Olin-Scheller, 2013). Despite out-of-school learning outcomes, some students may underachieve in the classroom as informal learning does not automatically transfer to language skills tested formally. To show students that all types of language learning are valuable, teachers could, at times, consider measuring vocabulary knowledge using freely available vocabulary tests such as DIALANG lexical proficiency tests (<https://dialangweb.lancaster.ac.uk/>). This would align with the principles of the *Common European Framework of Reference for Languages* (2001), which encourage teachers to appreciate and document the whole range of learners' language skills, whether attained formally or informally. Receiving feedback on vocabulary acquired both in-school and out-of-school might bridge L2 English learners' motivational gap between formal and informal learning and make learners realize that different learning environments complement each other.

Appendix

Correlations between the frequency and the amount of English in online activities²

Music

1st year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 26	$r = .191, p = .205$	$r = .167, p = .267$	$r = .323, p = .028$
<i>Less English Daily</i> n = 12	$r = -.160, p = .289$	$r = -.134, p = .375$	$r = -.205, p = .172$
<i>Less English Weekly/ English Rarely*</i> n = 8	$r = -.014, p = .928$	$r = -.008, p = .958$	$r = -.187, p = .212$
2nd year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 25	$r = .200, p = .182$	$r = .270, p = .069$	$r = .211, p = .160$
<i>Less English Daily</i> n = 11	$r = -.328, p = .026$	$r = -.290, p = .050$	$r = -.192, p = .202$
<i>All-English Weekly/ Less English Weekly/ English* Rarely</i> n = 10	$r = .97, p = .522$	$r = -.026, p = .863$	$r = -.056, p = .712$

Films

1st year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 24	$r = .518, p = .000$	$r = .485, p = .001$	$r = .476, p = .001$
<i>Less English Daily</i> n = 6	$r = -.219, p = .14$	$r = -.173, p = .250$	$r = -.258, p = .084$
<i>All-English Weekly/ Less English Weekly*</i> n = 11	$r = -.294, p = .048$	$r = -.290, p = .051$	$r = -.161, p = .286$
<i>English Rarely</i> n = 5	$r = -.192, p = .201$	$r = -.194, p = .197$	$r = -.266, p = .074$
2nd year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 22	$r = .371, p = .011$	$r = .451, p = .002$	$r = .351, p = .017$
<i>Less English Daily</i> n = 6	$r = -.326, p = .027$	$r = -.316, p = .032$	$r = -.241, p = .107$
<i>All-English Weekly/Less English Weekly*</i> n = 11	$r = -.294, p = .047$	$r = -.322, p = .029$	$r = -.255, p = .087$
<i>English Rarely</i> n = 7	$r = .140, p = .355$	$r = .051, p = .736$	$r = .041, p = .787$

2. Categories* including only a few participants were combined for analysis (cf., Table 1).

Sites

1st year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 19	$r = .548, p = .000$	$r = .497, p = .000$	$r = .552, p = .000$
<i>Less English Daily</i> n = 9	$r = .114, p = .450$	$r = .043, p = .777$	$r = -.030, p = .845$
<i>All-English Weekly/ Less English Weekly*</i> n = 9	$r = -.517, p = .000$	$r = -.423, p = .003$	$r = -.389, p = .007$
<i>English Rarely</i> n = 9	$r = -.275, p = .065$	$r = -.233, p = .119$	$r = -.310, p = .036$
2nd year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 21	$r = .536, p = .000$	$r = .524, p = .000$	$r = .421, p = .004$
<i>Less English Daily</i> n = 10	$r = -.275, p = .064$	$r = -.246, p = .099$	$r = -.207, p = .167$
<i>All-English Weekly/ Less English Weekly*</i> n = 8	$r = -.330, p = .025$	$r = -.200, p = .182$	$r = -.044, p = .769$
<i>English Rarely</i> n = 7	$r = -.078, p = .606$	$r = -.232, p = .122$	$r = -.299, p = .044$

Games

1st year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily/ Weekly*</i> n = 25	$r = .691, p = .000$	$r = .667, p = .000$	$r = .560, p = .000$
<i>Less English Weekly/ English Rarely*</i> n = 21	$r = -.691, p = .000$	$r = -.667, p = .000$	$r = -.560, p = .000$
2nd year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily/ Weekly*</i> n = 25	$r = .514, p = .000$	$r = .571, p = .000$	$r = .390, p = .007$
<i>Less English Weekly/ English Rarely*</i> n = 21	$r = -.489, p = .000$	$r = -.509, p = .000$	$r = -.404, p = .005$

Games+

1st year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 11	$r = .520, p = .000$	$r = .554, p = .000$	$r = .558, p = .000$
2nd year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 11	$r = .355, p = .015$	$r = .468, p = .001$	$r = .596, p = .000$

Social Network Sites

1st year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 7	$r = .140, p = .353$	$r = .067, p = .656$	$r = .058, p = .708$
<i>Less English Daily</i> n = 24	$r = -.303, p = .041$	$r = -.299, p = .043$	$r = -.205, p = .171$
<i>All-English/Less English Weekly/ English Rarely*</i> n = 15	$r = .196, p = .193$	$r = .215, p = .151$	$r = .146, p = .332$
2nd year	High-frequency	Mid-frequency	Low-frequency
<i>All-English Daily</i> n = 7	$r = .179, p = .234$	$r = .155, p = .303$	$r = .064, p = .674$
<i>Less English Daily</i> n = 15	$r = -.203, p = .177$	$r = -.225, p = .132$	$r = -.138, p = .362$
<i>All-English/Less English Weekly*</i> n = 12	$r = -.051, p = .737$	$r = -.040, p = .789$	$r = -.079, p = .601$
<i>English Rarely</i> n = 12	$r = .121, p = .424$	$r = .154, p = .306$	$r = .174, p = .248$

References

- Bocconi, S., Chiocciariello, A., & Earp, J. (2018). The Nordic approach to introducing computational thinking and programming in compulsory education. Report prepared for the Nordic@BETT2018 Steering Group. <https://doi.org/10.17471/54007>
- Brevik, L. M. (2019). Gamers, surfers, social media users: Unpacking the role of interest in English. *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.12362>
- Brevik, L. M. (2016). The Gaming Outliers: Does out-of-school gaming improve boy's reading skills in English as a second language? In E. Elstad, (Ed.), *Educational Technology and Polycontextual Bridging* (pp. 39–61), Rotterdam, the Netherlands: Sense Publishers.
- Cabot, M. (2018). Personal English Learning Ecologies and Meaningful Input with Digital and Non-Digital Artefacts. *Nordic Journal of Digital Literacy*, 13(2), 94–112. <https://doi.org/10.18261/issn.1891-943x-2018-02-03>
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning and Technology*, 11(3), 38–63.
- Cobb, T. (2016). Numbers or numerology? A response to Nation (2014) and McQuillan (2016). *Reading in a Foreign Language*, 28(2), 299–304.
- Council of Europe (2018). *Common European Framework of Reference for languages: Learning, Teaching, Assessment. Companion volume with new descriptors*. Language Policy Programme. Education Policy Division Education Department. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>. Accessed 11.5.2018.
- Council of Europe (2001). *Common European Framework of Reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Coxhead, A., & Bytheway, J. (2014). Learning vocabulary using two massive online resources: you will not blink. In D. Nunan & J. Richards (Eds.), *Language Learning beyond the Classroom* (pp. 65–74). New York, London: Routledge.

- Dörnyei, Z., & Chan, L. (2013). Motivation and vision: an analysis of future L2 self images, sensory styles, and imagery capacity across two target languages. *Language Learning*, 63(3), 437–462. <https://doi.org/10.1111/lang.12005>
- Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental vocabulary acquisition through reading. *Language Teaching Research*, 16(2), 227–252. <https://doi.org/10.1177/1362168811431377>
- Education First (2017). EF EPI. English Proficiency Index. Retrieved from <http://www.ef-epi-2017-english.pdf>. Accessed 26.8.2019.
- Elgort, I., & P. Warren (2014). L2 vocabulary learning from reading: explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning*, 64(2), 365–414. <https://doi.org/10.1111/lang.12052>
- Eurostat. Being young in Europe (2015). https://ec.europa.eu/eurostat/statistics-explained/index.php/Being_young_in_Europe_today_-_digital_world. Accessed 15.11.2018.
- Finnish National Agency for Education (EDUFI). <https://www.oph.fi/english>. Accessed 19.4.2018.
- Gass, S. (1999). Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21(2), 319–333. <https://doi.org/10.1017/S0272263199002090>
- Gill-Flores, J., Torres-Gordillo, J.-J., & Perera-Rodríguez, V.-H. (2012). The role of online reader experience in explaining students' performance in digital reading. *Computers & Education*, 59, 653–660. <https://doi.org/10.1016/j.compedu.2012.03.014>
- Godwin-Jones, R. (2019). Riding the digital wilds: Learner autonomy and informal language learning. *Language Learning & Technology*, 23(1), 8–25. <https://doi.org/10.125/44667>
- González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? The effects of frequency and amount of exposure. *ITL International Journal of Applied Linguistics*, 166(1), 94–126. <https://doi.org/10.1075/itl.166.1.03fer>
- Hannibal-Jensen, S. (2017). Gaming as an English language learning resource among young children in Denmark. *CALICO Journal*, 34(1), 1–19. <https://doi.org/10.1558/cj.29519>
- Härmälä, M., Huhtanen, M., & Puukko, M. (2014). *Englannin kielen A-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013* [Learning outcomes of English A-course at the end of the comprehensive school]. Helsinki: Kansallinen koulutuksen arviointikeskus. English summary, pp. 11–12.
- Jalkanen, J., & Vaarala, H. (2013). Digital texts for learning Finnish: shared resources and emerging practices. *Language Learning and Technology*, 17(1), 107–124.
- Kerekes, E. (2014). Using songs and lyrics in out-of-class learning. In D. Nunan & J. Richards (Eds.), *Language Learning beyond the Classroom* (pp. 33–42). New York, London: Routledge.
- Krashen, S.D. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.
- Kumpulainen, K., & Mikkola, A. (2016). Toward hybrid learning. In E. Elstad, (Ed.), *Educational Technology and Polycontextual Bridging* (pp. 15–37), Rotterdam, the Netherlands: Sense Publishers.
- Kuppens, A. 2010. Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, 35(1), 65–85. <https://doi.org/10.1080/17439880903561876>
- Lan, Y., Chen, N., Ping, L., & Grant, S. (2015). Embodied cognition and language learning in virtual environments. *Education Technology Research and Development*, 63, 639–644. <https://doi.org/10.1007/s11423-015-9401-x>
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1–26. <https://doi.org/10.1093/applin/22.1.1>
- Laufer, B., & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30.
- Lee, S., & Pulido, D. (2017). The impact of topic interest, L2 proficiency, and gender on EFL incidental vocabulary acquisition through reading. *Language Teaching Research*, 21(1), 118–135. DOI: 10.1177/1362168816637381. <https://doi.org/10.1177%2F1362168816637381>
- Lin, P. (2014). Investigating the validity of the internet television as a resource for acquiring L2 formulaic sequences. *System*, 42(1), 164–176. <https://doi.org/10.1016/j.system.2013.11.010>
- Lin, P., & Siyanova-Chanturia, A. (2014). Internet television for L2 vocabulary learning. In D. Nunan & J. Richards (Eds.), *Language Learning beyond the Classroom* (pp. 149–158). New York, London: Routledge.

- Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, 10(1), 105–129. <https://doi.org/10.1080/14790718.2012.679275>
- Mackey, A. & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum.
- McQuillan, J., & Krashen, S.D. (2008). Commentary: can free reading take you all the way? A response to Cobb. *Language Learning and Technology*, 12(1), 104–108.
- McQuillan, J. (2016). Time, texts, and teaching in vocabulary acquisition: A rebuttal to Cobb. *Reading in a Foreign Language*, 28(2), 307–318.
- Nation, P. (2016). Response to Tom Cobb. *Reading in a Foreign Language*, 28(2), 305–306.
- Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, 27(1), 136–145.
- Nation, P. (2014). How much input do you need to learn the most frequent 9000 words? *Reading in a Foreign Language*, 26(2), 1–16.
- Nation, P. (2006). How large a vocabulary is needed for reading and listening. *The Canadian Modern Language Review*, 63(1), 59–82. <https://doi.org/10.3138/cmlr.63.1.59>
- Nation, P. (1983). Testing and teaching vocabulary. *Guidelines*, 5, 12–25.
- OECD (2015). *Students, Computers and Learning: Making the Connection*, PISA, OECD Publishing. <https://doi.org/10.1787/19963777>
- Olsson, E., & Sylvén, L.K. (2015). Extramural English and academic vocabulary. A longitudinal study of CLIL and non-CLIL students in Sweden. *Apples – Journal of Applied Language Studies* 9(2), 77–103. <https://doi.org/10.17011/apples/urn.201512234129>
- Peer, W. v., Hakemulder, J. & Zyngier, S. (2012). *Scientific methods for the humanities*. Amsterdam, Netherlands; /Philadelphia, PA: John Benjamins Pub.
- Pellicer-Sanchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a foreign language*, 2(1), 31–55.
- Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL International Journal of Applied Linguistics* 169(2), 142–167. <https://doi.org/10.1075/itl.00010.pet>
- Peters, E.; Noreillie, A-S., Heylen, K., Bulté, B., & Desmet, P. (2019). The impact of instruction and out-of-school exposure to FL input on learners' vocabulary knowledge in two languages. *Language Learning* 69(3). <https://doi.org/10.1111/lang.12351>
- Peters, E. & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 40(3). Advance online publication. <https://doi.org/10.1017/s0272263117000407>
- Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, 63, 134–148. <https://doi.org/10.1016/j.system.2016.10.002>
- Rankin, Y., Gold, R., & Gooch, B. (2006). 3D role-playing games as language learning tools. A paper presented at the EuroGraphics 2006, September 4–8, 2006, Vienna, Austria http://www.thegooch.org/Publications/PDFs/Rankin_Gold_Gooch.pdf. Accessed 27.12.2017.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129>
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363. <https://doi.org/10.1177%2F1362168808089921>
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43. <https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503. <https://doi.org/10.1017/S0261444812000018>

- Sockett, G., & Toffoli, D. (2012). Beyond learner autonomy: a dynamic systems view of the informal learning of English in virtual online communities. *Procedia. Social and Behavioral Sciences*, 34, 212–215. <https://doi.org/10.1016/j.sbspro.2012.02.042>
- Statista, 2016. Statistics Finland. (n.d.). Internet usage for listening to music in Finland in Q1 2016, by age group. In *Statista – The Statistics Portal*. Retrieved March 23, 2017, from <https://www.statista.com/statistics/552155/internet-usage-for-listening-to-or-downloading-music-by-age-group/>. Accessed 23.3. 2017.
- Statista, 2015. Statistics Finland. (n.d.). Internet usage for searching for information on Wikipedia or other online encyclopedias in Finland in Q1 2015, by age group. In *Statista – The Statistics Portal*. <https://www.statista.com/statistics/551756/internet-usage-for-searching-for-information-by-age-group/>. Accessed 23.3. 2017.
- Sundqvist, P. (2009). *Extramural English Matters: Out-of-school. English and its Impact on Swedish Ninth Graders' oral proficiency and vocabulary*. Karlstad: Karlstad University Studies.
- Sundqvist, P. (2019). Commercial-off-the-shelf games in the digital wild and the L2 learner vocabulary. *Language Learning & Technology*, 23(1), 87–113. <https://doi.org/10.125/44674>
- Sundqvist, P., & Olin-Scheller, C. (2013). Classroom vs. extramural English: teachers dealing with demotivation. *Language and Linguistics Compass*, 7(6), 329–338. <https://doi.org/10.1111/lnc3.12031>
- Sundqvist, P., & Wikström, P. (2015). Out-of-school digital gameplay and in-school L2 English vocabulary outcomes. *System*, 51, 65–76. <https://doi.org/10.1016/j.system.2015.04.001>
- Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.) *Input in Second Language Acquisition* (pp. 235–253). Rowley MA: Newbury House.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 19–30). Cambridge: Cambridge University Press.
- Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL*, 24(3), 302–321. <https://doi.org/10.1017/S095834401200016X>
- Toffoli, D. & Sockett, G. (2010). How non-specialist students of English practice informal learning using Web 2.0 tools. *ASP: La revue du GERAS*, 58, 125–144. <https://doi.org/10.4000/asp.1851>
- Webb, S., & Rodgers, M.P.H. (2009a). The vocabulary demands of television programs. *Language Learning*, 59(2), 335–366. <https://doi.org/10.1111/j.1467-9922.2009.00509.x>
- Webb, S. & Rodgers, M.P.H. (2009b). The lexical coverage of movies. *Applied Linguistics* 30(3), 407–427. <https://doi.org/10.1093/applin/amp010>
- Vandergriff, I. (2016). *Second-language Discourse in the Digital World: Linguistic and Social Practices in and beyond the Networked Classroom*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Verspoor, M. H., De Bot, K., & van Rein, E. (2011). English as a foreign language: the role of out-of-school language input. In A. De Houwer & A. Wilton (Eds.), *English in Europe today: Sociocultural and Educational Perspectives* (pp. 147–166). Amsterdam; Philadelphia: John Benjamins Pub.Co.
- Voulgari, I., Komis, V., & Sampson, D.G. (2014). Learning outcomes and processes in massively multiplayer online games: exploring the perceptions of players. *Education Technology Research Development* 62, 245–270. DOI 10.1007/s11423-013-9312-7. Accessed 23.3. 2017.
- Zheng, D., Bischoff, M., & Gilliland, B. (2015). Vocabulary learning in massively multiplayer online games: context and action before words. *Education Technology Research Development* 63, 771–790. <https://doi.org/10.1007/s11423-015-9387-4>