English Vocabulary Learning Through Watching YouTube Video Blogs and Reading Blog Posts

Henriette L. Arndt

University of Oxford

Department of Education

St. Anne's College

Contact: henriette.arndt@education.ox.ac.uk

A thesis submitted in partial fulfilment of the requirements for the degree of *MSc Applied Linguistics and Second Language Acquisition*

August, 2014

Table of Contents

List of Tables and Figures ..... **iii**

Abstract ..... **iv**

Acknowledgements ..... **v**

Chapter 1: Introduction **1**
1.1. Social Media and Language Learning **1**
1.2. Virtual Communities **2**

Chapter 2: Selected Review of the Literature **4**
2.1. Online Informal Learning of English **4**
2.2. Learning Mechanisms in OILE **5**
2.2.1. Incidental and Intentional Learning 6
2.3. Assessing Vocabulary Knowledge: What Does it Mean to Know a Word? **8**
2.3.1. Aspects of Vocabulary Knowledge 9
2.3.2. Receptive and Productive knowledge 10
2.3.3. Partial and Precise Knowledge 10
2.4. Research on Incidental Vocabulary Learning **11**
2.4.1. Vocabulary Gains from Reading 13
2.4.2. Vocabulary Gains from Watching Videos 17
2.4.3. Studies Comparing Gains from Reading and Watching Videos 20
2.5. Variables That Affect Incidental Vocabulary Learning **23**
2.5.1. Coverage and Vocabulary Size 23
2.5.2. In-text Frequency 24

Chapter 3: Research Questions and Hypotheses **26**

Chapter 4: Methodology **27**
4.1. Design **27**
4.2. Participants **27**
4.3. Procedure **28**
4.3. Input Materials **29**
4.4. Target Words **31**
4.4.1. Word Forms, Repetitions, and Context 31
4.4.2. Pseudo-Words 32
4.4.3. Coverage and Comprehension 32
4.5. Dependent Measures and Scoring **33**
4.5.1. Orthography 35
4.5.2. Recall of Meaning 35
4.5.3. Recall of Grammatical Function 35
4.5.4. Recognition of Grammatical Function 36
4.5.5. Recognition of Meaning 36
4.6. Ethical Considerations **37**
4.6.1. Participant Information and Consent 37
4.6.2. Online Safety 38
4.6.3. Confidentiality and Anonymity 39

Chapter 5: Results **41**
5.1. Comparing the Experimental Groups **41**
5.2. Research Question 1: Total Gains **42**
5.3. Research Question 2: Aspects of Vocabulary Gains **44**
5.4. Research Question 3: Interactions with the Social Media **46**

5.5. Research Question 4: Appeal and Familiarity of Input Materials 46

Chapter 6: Discussion 51
6.1. Research Question 1: Total Gains 51
6.2. Research Question 2: Aspects of Vocabulary Gains 52
6.3. Generalisability of the Recorded Gains 53
6.4. Research Question 3: Interactions with the Input Media 56
6.4.1. User-user interaction 56
6.4.2. User-content interaction 57
6.5. Research Question 4: Appeal and Familiarity of Input Materials 59
6.6. Limitations and Suggestions for Future Research 61

Chapter 7: Conclusion 62

References 64

Appendices 73
Appendix 1: Input Materials: The three blog posts 73
1.1. Chris on Making YouTube friends 73
1.2. Lidewij on Online Friendships 74
1.3. JP on International Friendships 75
Appendix 2: Dependent Measures 77
2.1. Orthography 77
2.2. Recall of Meaning 78
2.3. Recall of Grammatical Function 79
2.4. Recognition of Grammatical Function 80
2.5. Recognition of Meaning 82
Appendix 3: Example of Participant Questionnaire (Video Group) 84
Appendix 4: Histograms Showing the Distribution of Gain Scores on Five Measures of Vocabulary Knowledge 85
4.1. Orthography 85
4.2. Grammatical Function (Recognition) 86
4.3. Grammatical Function (Recall) 87
4.4. Meaning (Recognition) 88
4.5. Meaning (Recall) 89

List of Tables and Figures

Table 1: Knowing a word (Nation, 1990: 31) 9
Table 2: Final group sizes 29
Table 3: Repetitions of pseudo-words 32
Table 4: Distribution of demographic variables and VLT scores 42
Table 5: Gains of five aspects of vocabulary knowledge 45

Figure 1: Distribution of total gains in the video group 43
Figure 2: Distribution of total gains in the blog group 43
Figure 3: Distribution of total gains from the two media 44
Figure 4: Did you like the videos you just watched/posts you just read? 47
Figure 5: How interesting did you find the topic of the videos/blog posts? 48
Figure 6: To what extent where the video blogs/blog posts like those that you usually watch/read? 48
Figure 7: How often do you read blogs/watch vlogs? 49
Figure 8: How many channels/blogs do you subscribe to? 50
Figure 9: Distribution of orthographic gains in the video group 85
Figure 10: Distribution of orthographic gains in the blog group 85
Figure 11: Distribution of gains in recognition of grammatical function in the video group 86
Figure 12: Distribution of gains in recognition of grammatical function in the blog group 86
Figure 13: Distribution of gains in recall of grammatical function in the video group 87
Figure 14: Distribution of gains in recall of grammatical function in the blog group 87
Figure 15: Distribution of gains in recognition of meaning in the video group 88
Figure 16: Distribution of gains in recognition of meaning in the blog group 88
Figure 17: Distribution of gains in recall of meaning in the video group 89
Figure 18: Distribution of gains in recall of meaning in the blog group 89

Abstract

Research on the *Online Informal Learning of English* has found that social media can facilitate language learning, mainly in the domain of vocabulary. It does so not only by providing the necessary language input, but also through the unique ways in which social media platforms encourage interaction between the user and the content as well as with other users. The present study investigated English vocabulary acquisition from engagement with two different types of social media: blog posts and video blogs. It also explored how learners make use of the unique opportunities for user-content and user-user interaction offered by these media. For this purpose, screen-captures of participants engaging with the blog posts and vlogs were collected.

The results showed that incidental vocabulary learning occurred, in approximately equal amounts, from reading blog posts and watching video blogs. They also revealed that different types of vocabulary knowledge were gained from the two input modalities. Written input was found to promote greater gains in orthographic knowledge than videos. In turn, the videos promoted greater recognition and recall of the target words' meaning and recall of their grammatical function. The analysis of the screen-captures indicated that learners did not often make use of the unique features offered by the blog and video platforms. The findings are discussed in the wider context of user engagement with social media and virtual communities.

Keywords: online language learning, informal learning, incidental vocabulary acquisition, online resources, blog, video blog,

Acknowledgements

I am sincerely grateful to my supervisor, Dr Robert Woore, for his insightful comments and constructive criticism as well as his encouragement, patience, and support throughout the duration of this project. I gratefully acknowledge the funding I received for this study from the Department of Education and St. Anne's College.

A special thank you also to the three videobloggers who have donated their time and creativity to produce the YouTube videos and blog posts that form the basis of this study. I would also like thank my classmates and friends for the illuminating discussions as well as the fun we have had this past year. Finally, I am deeply grateful to parents and brother for their unwavering support and reassurance. Without them I would not be where I am today.

Chapter 1: Introduction

1.1. Social Media and Language Learning

The term *social media* refers to websites and applications, which allow users to interact by creating and sharing multimedia content including text, images, and videos. They can take many forms, such as social networks, blogs, or image and video sharing portals.

On the one hand, educators seem to share a common belief that spending time on social media can lead to negative academic, personal, and social developments (Çetin, Sözcü, & Kinay, 2012; Toffoli & Sockett, 2013). On the other hand, engagement with social media seems to be more frequent, immediate, and relevant to learners than language tasks typically performed in classrooms or other formal learning contexts (Toffoli & Sockett, 2013). Second language acquisition (SLA) research has therefore recently sought to explore the potential of social media to support language learning. The main researchers in this field of *Online Informal Learning of English* are Geoffrey Sockett and Denyze Toffoli. Their studies show that language learning that can occur from engagement with online resources and that gains are primarily related to vocabulary acquisition, indicating that social media may be particularly useful in facilitating this type of learning (Sockett & Toffoli, 2012).

Inspired by this finding, the primary goal of the present study was to investigate English vocabulary acquisition from engagement with two different types of social media: blog posts and video blogs (henceforth *vlogs*). To the best of this author's knowledge, no study has yet explored the effectiveness of these materials for vocabulary learning. By directly comparing the amounts and types of lexical gains from these media, this study also addresses a gap in the existing research on second language (L2) vocabulary learning: while the field has produced numerous studies that investigate lexical learning from reading texts (e.g. Horst, Cobb, & Meara, 1998; Horst, 2005; Pigada & Schmitt, 2006; Waring & Takaki, 2003; Zahar, Cobb, & Spada, 2001) and watching videos (e.g. Hui, 2007; Markham, 1999; Rodgers, 2013; Winke, Gass, & Sydorenko, 2010; Yuksel & Tanriverdi, 2009), there is a distinct lack of studies that directly compare the

vocabulary gains from these two types of input (however, see Neuman & Koskinen, 1992; Vidal, 2011).

1.2. Virtual Communities

Toffoli and Sockett (2010) suggested that social media facilitate language learning not only by providing necessary language input, but also through the unique ways in which they encourage interaction between the user and the content and other users. The platforms investigated in the present study allow users to subscribe to their favourite content creators and leave comments on blog posts or videos. It can often be observed, that the availability of these features facilitates the formation of vibrant virtual communities around individual blogs or YouTube channels. Even though many language professionals may not be aware of such communities, their influence on current youth should not be underestimated. For example, a recent survey by the American entertainment-trade magazine *Variety* (Ault, 2014) among 1,500 readers aged 13 to 18 found that the five most popular celebrities among this age group were all YouTube stars, who far out-ranked celebrities from mainstream media.

One example of such a community are the *Nerdfighters*, from which participants in the present study were sampled. This community started forming in 2007 around the vlogs of American young-adult fiction author John Green and his brother, scientist and musician Hank Green ("vlogbrothers," 2014). In this community, things that are often labelled as 'nerdy' in a derogative fashion are highly valued and celebrated, for example intelligence and learning or being passionate and enthusiastic about your interests (whether an academic discipline, a book, a TV show, etc). Nerdfighters are primarily comprised of adolescents and young adults, but there are also older members (Green, 2014). Though largely based in the US, there are many members in all parts of the world (ibid.), who mainly network with each other online but often also attend gatherings and form close friendships in person. Most nerdfighters communicate in English although many are not native speakers and new members often rapidly improve their English proficiency when they start actively participating in the community. This anecdotal evidence was the initial rationale for the present study.

The secondary goal of the present study was therefore to explore how learners make use of the unique features offered by blogs and vlogs to engage with the content itself as well as with other users. For this purpose, screen-captures of participants engaging with the media were collected and analysed.

Chapter 2: Selected Review of the Literature

2.1. Online Informal Learning of English

A consequence of the ever-increasing access to the internet is that learners of English as a second or foreign language (ESL and EFL) have more frequent exposure to and interaction with contents authored by native speakers or other English language users. A portion of recent SLA literature has therefore been dedicated to exploring the potential of such content in supporting language learning, particularly in classroom settings (e.g. Blattner & Fiori, 2009; Blattner & Lomicka, 2012; Lomicka & Lord, 2009; Thorne & Reinhardt, 2008).

Sockett and Toffoli criticised the classroom-centeredness of this research, for implying that 'the potential for learning outside teacher-controlled or teacher orchestrated environments is [...] invalid' (Sockett & Toffoli, 2012: 139). However, they argued that this can occur, for example in the form of Online Informal Learning of English (OILE) when individuals, driven by the intention to communicate, engage with English-language media online, outside of a formal classroom or curriculum, with language learning occurring as a by-product of this engagement (Sockett, 2011 a.c.i. Toffoli & Sockett, 2013). In a recent series of studies, the two researchers have begun to explore the extent to which learners of English engage with online resources and the language learning that can result (e.g. Sockett & Toffoli, 2012; Sockett, 2011a, 2012b; Toffoli & Sockett, 2010).

First, Toffoli and Sockett (2010) conducted a large-scale survey of 222 French university-level EFL learners, asking which types of English-language online activities these students engaged in and how frequently they did so. 97% of the respondents stated that they regularly (weekly-monthly) engaged with English content online, mostly for entertainment or information-seeking purposes. Furthermore, participants reported that they spent most time watching films and television series or listening to music in English and slightly less time reading English texts, or communicating in English on social networks.

To learn more about the specific activities learners engage in, how much time they spend on them, and whether these behaviours are stable over time, Sockett and Toffoli (2012) followed up with a diary study involving five French university students. They were asked to record which resources they used and for how long, which language skills were involved in processing this material, and to

comment on the language they learned. As in the survey, the main activities reported by the participants in this study were, in order of frequency, watching films and television series, social networking, listening to music (sometimes while reading lyrics), and reading English-language content. The participants considered watching videos particularly helpful for developing their listening skills. However, most learners' comments actually related to new words or phrases that they encountered, indicating the occurrence of vocabulary learning.

Altogether, the studies reviewed above indicate that engagement with online materials can result in language acquisition, mainly in the domain of vocabulary. The present study builds on Toffoli and Sockett's work in that it explores in more detail the vocabulary acquisition from two types of online resources with which their participants reported engaging frequently: video materials and written text.

2.2. Learning Mechanisms in OILE

Toffoli and Sockett (2010: 4) describe *informal learning* as being 'unofficial, unscheduled, and impromptu', meaning that it does not happen in a classroom or formal learning context but is emergent from the learners' interactions with online resources. They suggested that these resources facilitate language learning by providing necessary language input, often in more than one modality (e.g. visual and auditory), and through the unique ways in which they encourage interaction between the user and the content and other users.

In this description of the mechanisms behind Online Informal Learning of English, Toffoli and Sockett (2010) draw on well-known SLA theories, which posit that L2 learning requires comprehensible language input (cf. Input Hypothesis, Krashen, 1981), the chance to produce language (cf. Output Hypothesis Swain, 1985), and meaningful interaction with others using the L2 (cf. Interaction Hypothesis Long, 1981). Woo and Reeves (2007) similarly claim that the main potential of web-based media to further language learning lies in its ability to facilitate meaningful interaction. Due to the limited scope of the present discussion, these theories cannot be discussed here in detail (however, see Gass, 1999b for an in-depth view). The notion that materials which provide input in more than one modality is helpful in facilitating learning draws on the theory of

Multi-Channel communication, which lies at the heart of multimedia approaches to education (discussed in depth in Moore, Burton, & Myers, 2004).

Because OILE is not planned by the learners (who may not even be aware of its occurrence), Sockett and Toffoli (2012) concede that their conception of informal learning shows parallels to *incidental learning*. As the incidental-intentional dichotomy is central to the vocabulary acquisition research upon which the present study is based, definitions and operationalisations of these two types of learning are discussed in detail in below.

2.2.1. Incidental and Intentional Learning

Firstly, some scholars use the term incidental learning to describe acquisition which occurs in the absence of a conscious *intention* to learn or rather when learners' attention is not focused on language learning, but on understanding the material they are engaging with (e.g. Gass, 1999a; Hulstijn, 1990; Laufer & Hulstijn, 2001; Nation, 2001; Schmitt, 2000). In this way, incidental learning has also been described as the by-product of communicative language tasks, such as processing written or spoken texts for their meaning, as opposed to being the target of these activities (e.g. Bruton, López, & Mesa, 2011; Huckin & Coady, 1999; Hulstijn, 2001). Conversely, intentional learning is thought to occur when learners' attention is focused on acquiring new language features, for example, vocabulary (Nation, 2001; Schmitt, 2000).

However, scholars widely accept that to acquire a new language feature learners must notice it in the input, which requires a certain degree of consciousness *(i.e. the Noticing Hypothesis,* Schmidt, 1990). Therefore, the difference between the two types of learning cannot lie only in the presence or lack of attention. The Depth of Processing Theory (Craik & Lockhart, 1972; Craik & Tulving, 1975) therefore suggests that the two types of learning differ *quantitatively,* not *qualitatively* with regard to the learner attention involved, so that the extent to which a learner consciously engages with a new word stands in direct relation to the depth of understanding she will obtain, leading in turn to more acquisition. Alternatively, R. Ellis (1999) proposed that the difference may lie in where learners place their 'focal' and 'peripheral' attention: in intentional learning, the focal attention is preoccupied with the linguistic forms themselves

(e.g. mapping form-meaning connections) and only peripheral attention is given to understanding the overall message, whereas in incidental learning the focal attention is on comprehension and only peripheral attention is dedicated to form.

All in all, it seems that a clear distinction between incidental and intentional learning is not easy to maintain on the basis of attention allocation alone. Furthermore, operationalising attention-based definitions poses a problem for research, because learners' minds cannot be directly accessed to determine whether or not they are focussing on form at any given point in time (Gass, 1999a).

Hulstijn (2001) therefore proposed to operationalise incidental vocabulary learning in research as occurring when learners engage with a text while not being forewarned that they would later be tested on vocabulary gains. His operationalisation has been widely accepted for its convenience. For example, without exception, the studies reviewed in section 2.4. have been classified as investigating incidental vocabulary learning on the same grounds: participants were reading a text or watching a video without knowing that their vocabulary acquisition would be tested.

Nevertheless, some aspects of Hulstijn's (2001) operationalisation remain problematic. Firstly, by assuming that the instructions to a learning task directly determine the type of learning will occur, it disregards what Gass (1999a: 320) described as 'the active role of the learner'. That is, even when not forewarned of an upcoming vocabulary test, a learner may, at any moment, choose to focus on learning new vocabulary instead of, or in addition to, comprehending the text's meaning. This happened, for example, in a study by Pellicer-Sánchez and Schmitt (2010), in which one participant underlined words in a text he was given to read for comprehension and reported trying to learn their meaning. Conversely, it is equally problematic to suppose that, when participants are forewarned of a vocabulary test, their main focus will be word learning and not meaning comprehension (Bruton et al., 2011). For example, in an interview study by Peters (2006), participants were told that they would be tested on vocabulary acquisition, yet still reported that their main focus was on text comprehension. In fact, whether or not a word is central to comprehending a text seems to be a more significant predictor of learning outcomes than knowledge of an upcoming test (Peters et al., 2009).

The review above emphasised the fact that learning which occurs when an individual is processing language to derive meaning (e.g. during OILE) can be classified both as incidental or intentional, depending on any individual learner's focus at any given moment. While the learner's main goal when engaging with online materials for entertainment or information-seeking purposes would be to understand their meaning, he may also choose to focus his attention on deriving and retaining the meaning of any unknown words he encounters.

Therefore, some scholars have recently proposed to abandon the incidental-intentional dichotomy altogether. Bruton et al. (2011), for example, argued that research should focus on the ways in which external conditions, such as the properties of the texts and tasks with which the learners are engaging, influence the amount and type of learning instead of on the learner-internal processes implied in the incidental-intentional dichotomy.

Following this suggestion, the main purpose of the present study is to compare vocabulary learning gains under two informal learning conditions: while reading blog posts and watching vlogs. The term incidental vocabulary will be used throughout this dissertation to describe lexical learning from engaging with L2 materials where learners are not forewarned of an upcoming vocabulary test. No claims are made, however, about whether the participants in this study were intentionally focusing on learning the target words, or whether the observed gains occurred as a by-product of purely meaning-focused processing.

2.3. Assessing Vocabulary Knowledge: What Does it Mean to Know a Word?

Before discussing previous research on incidental vocabulary acquisition, it is important to define the goal of vocabulary learning, i.e. what it means to know a word. Most studies have only measured learners' ability to understand the meaning of a word or produce forms for given meanings (e.g. Laufer & Nation, 1999; Nation, 1983; Schmitt, Schmitt, & Clapham, 2001). Most researchers would agree, however, that lexical knowledge is complex and comprises more than understanding which meanings correspond to which forms, wherefore a combination of tasks should be used to evaluate a variety of types of knowledge (Henriksen, 1999). The following sections give an overview of three commonly

distinguished dimensions of vocabulary knowledge, which were considered in the assessment of vocabulary gains in the present study.

2.3.1. Aspects of Vocabulary Knowledge

SLA researchers widely agree that knowing a word involves a number of *aspects* of knowledge. Probably the most well-known framework of vocabulary knowledge, proposed by Nation (1990, 2001; Table 1), distinguishes eight aspects as well as receptive and productive knowledge (section 2.3.2.). Many vocabulary researchers have since strived to include tests in their work, which measure different aspects of lexical knowledge.

Form		
Spoken form	R	What does the word sound like?
	P	How is the word pronounced?
Written form	R	What does the word look like?
	P	How is the word written and spelled?
Position		
Grammatical patterns	R	In what patterns does the word occur?
	P	In what patterns must we use the word?
Collocations	R	What words or types of words can be expected before or after the word?
	P	What words or types of words must we use with this word?
Function		
Frequency	R	How common is the word?
	P	How often should the word be used?
Appropriateness	R	Where would we expect to meet this word?
	P	Where can this word be used?
Meaning		
Concept	R	What does the word mean?
	P	What word should be used to express this meaning?
Associations	R	What other words does this word make us think of?
	P	What other words could we use instead of this one?

R=receptive, P=productive

Table 1: Knowing a word (Nation, 1990: 31)

The present study included tests (adapted from Webb, 2007) of three of Nation's aspects: written form, grammatical pattern, and concept.

2.3.2. Receptive and Productive knowledge

Another commonly discussed dimension of vocabulary knowledge is the distinction between *receptive* and *productive knowledge*. Receptive knowledge has been defined as being able to understand various aspects of a word (form, position, function, and meaning, cf. Table 1) when encountering it in language input and productive knowledge as being able to retrieve the appropriate word (its form, syntactic properties, etc.) when trying to express a meaning in production (e.g. DeKeyser & Sokalski, 1996; Nation, 1990, 2001). It has been questioned, however, whether receptive and productive knowledge truly represent a dichotomy even though the terms are widely used in the SLA literature in this way. For example, Melka (1997) and Paribakht & Wesche (1993) suggested that the constructs should be seen instead as two ends of a knowledge continuum.

Dichotomy or continuum, productive knowledge as defined above is difficult and time-consuming to measure, as this requires recording and analysing learner language. Nevertheless, some vocabulary acquisition research has taken this approach: Neuman and Koskinen (1992), for example, counted the occurrences of target words in students' homework assignments after they had been exposed to these words through watching videos or reading texts, and Sockett (2012a a.c.i. Toffoli & Sockett, 2013) examined students' vocabulary uptake from watching TV series by analysing their writing (fan fiction).

However, most vocabulary learning studies (see section 2.4.) have taken a different approach to measuring receptive and productive lexical gains: operationalising receptive knowledge as being able to *recognise* an aspect of vocabulary knowledge (e.g. orthographic form or meaning) in a set of options and productive knowledge as being able to *recall* these aspects when prompted (e.g. recalling meaning when prompted with an orthographic form). The present study adopted this operationalisation of receptive and productive knowledge as recognition and recall.

2.3.3. Partial and Precise Knowledge

Finally, *partiality* and *precision* form a third dimension of vocabulary knowledge, which distinguishes between words that are only vaguely familiar to a learner and those of which she has full knowledge (Henriksen, 1999). Harley

(1995: 3), for example, pointed out that learning a word is not an all-or-nothing phenomenon: At first a learner may only recognise the form but not know its associated meaning, but after passing through phases of 'partial word knowledge' he might reach precise knowledge of all its aspects.

This dimension is closely related to the recognition-recall contrast: Research has shown that receptive knowledge usually precedes production (e.g. N. Ellis & Beaton, 1993; Laufer & Paribakht, 1998; Laufer, 1998). Norman (1976: 135) hypothesised that this is because recall requires more complete knowledge, while recognition 'require[s] only sufficient [knowledge] about a word to distinguish it from all other possibilities'. Norman's claim can also explain the common finding that the size of a learner's receptive vocabulary is larger than the productive one (e.g. N. Ellis & Beaton, 1993; Laufer & Paribakht, 1998; Laufer, 1998). Research on vocabulary acquisition should thus aim to include tests, which allow learners to demonstrate partial as well as precise knowledge of the lexical items in question, as was the case in the present study.

Throughout the review above it was repeatedly emphasised that studies of vocabulary acquisition should include multiple tests to measure a variety of dimensions of lexical knowledge. The present study therefore assessed both recognition and recall of the target words' written forms, grammatical behaviour, and conceptual meaning. Furthermore, the tests were designed and scored in a way that credited partial as well as precise knowledge.

2.4. Research on Incidental Vocabulary Learning

In first language (L1) acquisition research, the suggestion that vocabulary can be acquired through reading has commonly been used to account for the large vocabularies that children develop without formal instruction as well as the acquisition of low-frequency words, which occur infrequently in everyday conversation (e.g. Anderson, Wilson, & Fielding, 1988; Nagy & Herman, 1987). SLA scholars quickly adopted this theory and began to investigate vocabulary learning through reading L2 materials.

Seminal to this branch of research is Saragi, Nation and Meister's (1978) study, in which they tested native English speakers' understanding of *nadsat* words (Russian slang) after having read Anthony Burgess's novella *A Clockwork*

Orange (published in 1962). It was assumed that the participants had not previously known these words and a surprise multiple-choice test of meaning recognition found that most were later able to recall the meaning of a majority of them (69 out of 90; 77%). Saragi et al.'s (1978) study has been criticised, however, for not accurately representing L2 vocabulary learning. As Horst, Cobb and Meara (1998) pointed out, it is presumably more difficult to infer the meaning of unfamiliar L2 words in an L2, rather than L1, context. This and the possibility that the English context in Burgess's novel may have increased participants' attention to the Russian words could, therein have artificially increased the vocabulary gains in Saragi et al.'s (1978) study. Indeed, participants in later studies of incidental vocabulary learning have been found to learn proportionally fewer words from reading L2-only texts (see section 2.4.1.).

Nevertheless, Saragi et al. (1978) inspired further incidental vocabulary learning research. In empirical studies, participants are typically asked to engage with L2 input, usually a written text. They are not told that they will be tested on their vocabulary learning. Lexical gains are then assessed in a variety of ways, e.g. using multiple-choice, prompted recall, or translation tests. Measures are usually taken to ensure that the target words were previously unknown to the participants. This may be ascertained through pre-testing, or by replacing the target items with pseudo-words (e.g. Waring & Takaki, 2003; Webb, 2007, 2008).

Numerous studies have investigated incidental vocabulary acquisition from engagement with a variety of L2 media. Most have focused on assessing vocabulary gains through reading (e.g. Horst et al., 1998; Horst, 2005; Kweon & Kim, 2008; Pigada & Schmitt, 2006; Saragi et al., 1978; Waring & Takaki, 2003; Webb, 2007, 2008; Zahar et al., 2001), but video materials have also been studied frequently (e.g. Hui, 2007; Markham, 1999; Rodgers, 2013; Vidal, 2003, 2011; Winke et al., 2010; Yuksel & Tanriverdi, 2009). Other studies have been concerned with, for example, incidental vocabulary learning from listening to songs or reading comic books (Milton, 2008), playing online role-playing games (Huang & Yang, 2012) or social network-based flash games (Çetin et al., 2012).

In the following sections, some findings regarding vocabulary gains from written and audio-visual input will be reviewed in more detail, as they form the basis for the present study's rationale and design. The research presented below makes clear that incidental vocabulary acquisition is an incremental process, in which learners slowly build up knowledge of a new word through repeated encounters over time. Vocabulary gains are furthermore strongly influenced by a number of variables, which enable learners to acquire more new words in some contexts than in others.

2.4.1. Vocabulary Gains from Reading

Horst, Cobb, and Meara (1998) conducted what is probably the most well-known study of L2 incidental vocabulary learning through reading. The participants were 34 Arabic University students, whom the authors categorised as low-intermediate EFL learners. The participants listened to, and read along with, a graded reader over ten class periods to ensure that they would read the entire book and to preclude intentional word learning by focusing their attention on the story. Because the authors set out specifically to test the effects of the target words' general frequency and their frequency of occurrence within the text, the 45 target items were selected from a variety of frequency bands and occurred in the text between 2 and 17 times each. Participants' lexical gains were measured using two immediate pre- and post-tests: one multiple-choice test assessing recognition of meaning and one word association task testing form-meaning associations by asking students to identify *the odd one out*. Furthermore, to investigate the effect of the students' vocabulary size on their learning gains, Nation's (1990) Vocabulary Levels Test was administered together with the pre-test. On the multiple-choice test, participants showed mean gains of 4.6 (*SD* = 4.1) or 22% of previously unknown target words. Additionally, they made on average 1.3 (*SD* = 2.3) or 16% more correct word associations on the post-test than the pre-test. Significant correlations were found between learning gains and vocabulary size (r = .31), the words' general frequency (r = .14) and their in-text frequency (r = .49).

Zahar, Cobb, and Spada (2001) investigated vocabulary gains from reading in a study with 144 French-Canadian High School ESL learners. The students were selected from five levels of proficiency (intermediate to bilingual) as determined

by the school's placement exam. During class-time, participants read along while listening to a short story from an intermediate graded reader and were then given the rest of the lesson to re-read the story on their own. Knowledge gains were assessed through pre- and post-tests, administered thirteen days before and two days after the reading task. They required students to match a list of 30 target words to their definitions (recognition of meaning). On average, students learned 9% of the 30 words tested. Higher and lower proficiency participants exhibited fewer gains (5% and 6% respectively) than mid-level students (7-10%). The authors remark that the higher-level students' low gains can be attributed to a ceiling effect on the pre-test. They learned, on average, 1.5 of the 3.5 words (43%) that had been unknown to them while the lower-level students only learned 1.8 out of 19.2 (9%) previously unknown items. This finding emphasises the importance of assessing individual gains in relation to the available learning opportunities. Furthermore, the study found a correlation between the number of repetitions of target words in the text and variations in learning gains (r = .36) and a significant effect of the amount of information about a target word's meaning provided by the context in which the words appeared (contextual informativeness).

Waring and Takaki (2003) investigated how a variety of aspects of vocabulary knowledge were gained by L2 reading and how these gains were retained over time. Their study involved 15 lower intermediate ESL learners at a Japanese University. 25 words with varying in-text frequencies were selected from a graded reader and replaced with pseudo-words. There was no pre-test because the pseudo-words were unknown to the participants. They read the book and were tested on three aspects of vocabulary knowledge: word form recognition, meaning recognition, and meaning recall (L2-L1 translation). The post-test was administered three times: immediately, seven to ten days later, and three months later. Learning gains varied considerably depending on the type of test used to assess them: on the immediate post-test, the learners recognised an average of 15.3 word forms (61%, *SD* = 13%), 10.6 word meanings (42%, *SD* = 16%) and translated 4.6 words correctly (18%, *SD* = 8%). Of their original recognition gains, participants retained about 75% until the first and 56% until the second delayed post-test. Recall gains were less well retained (41% and 20%). Waring and Takaki's (2003) findings underline the importance of assessing

multiple aspects of vocabulary knowledge as well as using delayed post-tests when investigating incidental learning gains to assess their durability.

Pigada and Schmitt (2006) also found that gains varied depending on the type of vocabulary knowledge being assessed. With the aim of researching how extensive reading can increase knowledge of a large number of target words, they conducted a case study of one adult Greek beginner learner of French as a foreign language who, over the course of one month, read four graded readers (30,000 words in total). 133 target words appeared in the text from 1 to over 20 times and the participant was pre- and post-tested in interviews on recall of these words' meanings, orthographic forms, and grammatical behaviour (indicating suitable determiners for nouns and prepositions for verbs). The participant showed gains in orthographic knowledge of 24%, the meaning of 13%, and the grammatical function of 19% of the words previously unknown to him. Additionally, an analysis by word class revealed that he learned more of the spelling and meaning of verbs but gained more knowledge of the grammatical function of nouns. Pigada and Schmitt (2006) argued that the gains they recorded were more representative of long-term L2 vocabulary learning through reading. However, their study did not include delayed post-tests to assess the long-term durability of the reported gains and, because it was a case study, the generalisability is limited.

Horst (2005) assessed the learning from extensive reading in a bigger sample. He employed an innovative research design, creating individualised vocabulary tests for each participant to assess incidental vocabulary gains from reading materials that they themselves had chosen. Like Sockett and Toffoli (2012), Horst argued that giving learners the opportunity to choose which L2 materials to engage with would increase motivation, and therein learning gains. Horst's participants were 21 adult immigrant ESL learners, each of whom borrowed, and presumably read, on average ten to eleven graded readers from the project's library over the course of the six-week study. To assess prior knowledge, twelve graded readers were first scanned electronically to create lists of all words that appeared in the texts. A random sample was taken from these to build a 'baseline measure of knowledge of words that typically occur in graded readers' (Horst, 2005: 363). To test knowledge gains, samples from the most popular graded readers (covering approximately 67% of the materials

participants read) were used to create individualised post-tests for all participants with 50 items that occurred in their books. The target items were low-frequency words, which the author deemed unlikely to have been met by the learners outside of the readings. In the post-test, participants self-reported their knowledge of each target word using a simplified version of Paribakht and Wesche's (1996) Vocabulary Knowledge Scale (VKS). They indicated that they either knew or didn't know what a word meant, or had an idea but were unsure its exact meaning. Horst (2000 a.c.i. Horst, 2005) reported that this was a reasonably reliable measure: adult learners were able to provide translations of about 80% of words they rated as known.

In Horst's (2005) study, participants reported knowing 10.3 (*SD* = 7.6) or 50% more words on the post- than the pre-test. However, these results must be treated with caution as they are based on two different tests: the individualised post-test and the baseline measure. Horst therefore also asked participants to rate on the VKS some words they had identified as unknown on the baseline test and which also occurred in their readings. This measure provided similar results, indicating, on average, partial or full growth in 18 out of 35 words (ca. 50%).

All in all, the research reviewed in this section shows that incidental vocabulary learning does occur from L2 reading. However, the studies report a wide range of results, i.e. gains in knowledge of as little as 10% (Zahar et al., 2001) to as much as 61% (Waring & Takaki, 2003) of previously unknown words. The exact scores seem to depend on the type of test used to assess vocabulary knowledge as well as a number of learner- and text-based variables, such as vocabulary size, task demands, the number of encounters with the target words and the context surrounding them. Nevertheless, all of these studies support the conclusion that single incidental encounters do not lead to effective vocabulary learning but that, instead, repeated exposure to new lexical items is necessary.

There is some evidence that extensive and sustained input, e.g. in extensive reading programs, can lead to substantial cumulative gains, though more so in receptive than productive vocabulary (e.g. Elley, 1991; Horst, 2005; Kweon & Kim, 2008; Pigada & Schmitt, 2006). However, it is often suspected that many learners do not engage with L2 reading materials to this extent (see e.g. Waring, 2009). A number of studies have therefore explored incidental

vocabulary acquisition from engaging with other L2 materials. The following section reviews studies, which have assessed incidental vocabulary learning from watching videos, the second medium under investigation in the present study.

2.4.2. Vocabulary Gains from Watching Videos

Most studies of incidental vocabulary learning from videos have been aimed at comparing the effects of viewing them with or without captions. One of the first studies on the topic was conducted by Markham (1999). In a two-by-two design, Markham compared advanced ESL learners' recognition of written and spoken word forms after viewing one of two 12-13 minute documentaries (about American civil rights or whales) with or without captions. Each of 100 target words was repeated on average three times per video. Overall, the participants gained significantly more knowledge from watching the civil rights video than the whale documentary; the author hypothesized that this may be due to differences in comprehensibility, which was not measured. With both videos, however, the caption group significantly outperformed the non-caption group, recognizing an average of 72% (*SD* = 10.2%) and 67% (*SD* = 9.4%) of the target words versus 64% (*SD* = 11.7%) and 56% (*SD* = 12.1%) from the civil right and whale documentary respectively. As will be seen in the course of this discussion, these gains are comparatively large. On the one hand, this could be because the study did not include a pre-test, so that a proportion of the knowledge demonstrated on the word recognition tests may have existed before watching the videos. On the other hand, the high gains may be because only one aspect of lexical knowledge, form recognition, was assessed. Studies measuring multiple aspects of vocabulary learning usually find higher gains of knowledge of form than, for example, knowledge of meaning (e.g. Hui, 2007 below).

Hui (2007) tested the lexical gains of 187 Chinese EFL learners after watching a sixteen minute science documentary with L1 captions, L2 captions, or without captions. The sample contained a higher- and lower-proficiency group (based on University grade level: first year vs. third year undergraduate). To avoid priming participants towards the target words, a pre-test in the form of a binary self-report (I do/don't know this word) was conducted with a comparable control group. On the post-test, the participants were tested on their knowledge of ten target words that had been unknown to the control group. Three aspects of

their vocabulary knowledge were tested: recognition of phonological forms, spelling, and meaning. The gain scores varied considerably depending on which aspect of knowledge was assessed, the proficiency of the participants, and whether they watched the videos with no captions, L2 captions, or L1 subtitles. On average, participants gained knowledge of the spoken forms of 44% to 79% (composite *M* = 60%, *SD* = 10.7%)[1], the spelling of 35% to 68% (composite *M* = 49%, *SD* = 10.8%), and the meaning of 33% to 62% (composite *M* = 47%, *SD* = 10.7%) of the target words. The low proficiency participants who watched the videos without captions achieved the lowest scores. L2 captions, as compared to L1, significantly improved participants' performance on the tests of word form recognition (*d* = .63 and .74 for the low and high proficiency groups respectively)[2] and spelling (*d* = .98 and 1.2). However, on the test of meaning recognition, the L1 caption groups significantly outperformed the L2 caption groups (*d* = .63 and .36). The authors propose that this may be because students relied on L1 translation to decode the L2 meaning, which was aided by the L1 captions. The reliability of the numerical gains reported must, however, be questioned since they are based on different pre- and post-tests, conducted with different groups of learners and Hui (2007) does not specify how exactly the three aspects of word knowledge were tested.

In a study by Winke, Gass, and Sydorenko (2010), 150 participants viewed three animal documentaries, each three to five minutes in length, dubbed in their various L2s (Spanish, Russian, Arabic, and Chinese; L1 = English). All learners viewed each video two times; some saw the films without captions twice (*n* = 8), some saw them twice with L2 captions (*n* = 9), but most watched them once with and once without L2 subtitles (*n* = 133). Afterwards, the participants were asked if they had ever encountered the target words before seeing the videos (retrospective pre-test), and to translate them from their L2 to L1. In the translation test, each target word was presented twice, once orthographically and once aurally. All in all, the participants who had watched the videos with captions twice significantly outperformed both other groups, gaining on average

[1] The composite means and standard deviations were calculated by the author of this dissertation using the experimental groups' means and their standard deviations reported by Hui (2007).
[2] These effect sizes were calculated by the author of this dissertation using the means and standard deviations reported by Hui (2007)

knowledge of 16% (*SD* = 9.6%)[3] of the previously unknown words. Seeing the videos with captions once still significantly improved gains (composite *M* = 11%, *SD* = 11.0%) compared to not seeing captions at all (composite *M* = 3%, *SD* = 2.8%). Finally, participants who saw the film with captions first and then without significantly outperformed participants who saw the captions second. The authors furthermore report on the differential learning effects depending on the target L2; however, the discussion of these results would go beyond the scope of the present review. One limitation of the study is that the experimental groups contained widely different numbers of participants (the authors state no reason for this). Furthermore, the standard deviations were proportionally large and it must be questioned how reliable the gain scores are, since the test of previous vocabulary knowledge was a self-report measure administered *after* the treatment.

Finally, Rodgers (2013) tested the incidental vocabulary gains of 187 pre-intermediate to intermediate ESL learners, who watched ten episodes of the TV show *Chuck* over the course of eleven weeks. The study also included a control group with 86 participants, who received no treatment. The targets were 60 words with low general frequency (not in the most frequent 3,000 words), which occurred frequently (5 to 54 times) in the episodes. Gains were assessed using two pre- and post- multiple-choice meaning recognition tests, designed to tap into different levels of knowledge: the 'Tough Test', where the distractors were similar to the target words in either meaning or form, required more precise knowledge of the lexical items than the 'Sensitive Test', where the distractors and targets were unrelated. Participants in the experimental group demonstrated average gains of knowledge of 23% of previously unknown words on the Tough and of 30% on the Sensitive test (no standard deviations reported). The control group showed gains in the knowledge of 21% and 25% of previously unknown words on the two tests respectively. That is, both the gains of the target and control group were significant, as, however, were the differences between them. This indicates that, probably due to the extended length of the treatment, not all vocabulary gains can be attributed to the input but that watching the videos

[3] These numbers represent composite means and standard deviations of the participants' gains as demonstrated on the aural and written post-test combined. They were calculated by the author of this dissertation using the group means for both tests and their standard deviations as reported by Winke et al. (2010).

nevertheless had a small but significant effect. Rodgers (2013) also found that the target words' in-text frequency was significantly correlated with vocabulary gains (r = .30).

Rodgers' (2013) study is unique in that it investigates incidental vocabulary learning from extensive engagement with videos, namely multiple episodes of the same TV show. The gains seem low in comparison to short-term treatments, probably because the latter tend to be designed to repeat target words frequently in a short time. Corpus studies have, however, been conducted which show that themes, characters, and plots in films (Webb, 2010) or successive episodes of TV shows (Rodgers & Webb, 2011; Sockett, 2011a; Webb & Rodgers, 2009) also lead to repeated encounters with low frequency words, which lend themselves to incidental learning.

Altogether, the studies reviewed above indicate that incidental L2 vocabulary acquisition from watching videos does take place. Just like in research on incidental learning from reading, the reported gains vary considerably because they depend on a number of other variables. They range from gains of 3% (Hui, 2007) to 67% (Markham, 1999) of previously unknown words for uncaptioned viewing and from 16% (Hui, 2007) to 72% with captions (Markham, 1999).

2.4.3. Studies Comparing Gains from Reading and Watching Videos

Neuman and Koskinen (1992) conducted the first of only two studies to directly compare incidental vocabulary gains from reading and watching videos. 179 middle school ESL learners participated in this research. It is important to note that these children were selected from an *at risk* population, their academic performance being on average 2-3 years below their grade-level. Participants were divided into several groups, each of which was presented with a different type of input. Over the course of nine weeks, some watched nine segments of a science TV show, with or without captions, while others listened to and read along with the same segments' scripts. From each segment, a group of judges chose the ten most difficult lexical items (mainly academic vocabulary), which were used as target words in the incidental learning test.

The participants self-reported their knowledge of the target words (I do/don't know this word) at the beginning of the study. Their lexical gains were

then evaluated using a wide range of tests. Each week, students were tested on form recognition of that week's target words using a multiple-choice test with form-related distractors. The authors also counted the number of occurrences of the target words and the concepts behind them (idea units) in the students' weekly homework assignments. At three-week intervals, participants were given a test in which they were required to select, from four options, the sentences in which the target words were used in their correct senses (sentence anomaly test). Finally, at the end of the nine units, they completed a multiple-choice test assessing their ability to recognize the target words' meanings in context.

The captioned video group demonstrated significantly higher gains than the reading group on all measures (Cohen's *d* between 0.7 and 1.4)[4] and significantly better results than the no-caption group on all tests but one of the three sentence anomaly measures (Cohen's *d* between 0.5 and 0.7). The authors also found higher means for the no-caption group than for the reading group on almost all measures but did not report whether they are significantly different (however, sometimes $d > 0.2$, indicating small effects). Finally, Neuman and Koskinen (1992) reported that higher proficiency was associated with larger vocabulary gains and that contextual informativeness predicted which target words were more readily acquired.

All in all, Neuman and Koskinen's (1992) research indicates that video materials can serve as useful input for incidental vocabulary learning and that this type of learning can lead to greater gains in some aspects of vocabulary knowledge than engaging with reading materials. It is unclear, however, if the learning effects can be exclusively attributed to the engagement with the video segments/scripts because the study was conducted over the course of nine weeks (cf. Rodgers, 2013 above). Furthermore, the generalisability of this study's results to other L2 learner populations is questionable since the participants were selected because they had learning difficulties.

[4] To be able to summarize the results of the vast number of measures reported on in this study, the author of this dissertation calculated effect sizes using the means and standard deviations reported by Neuman and Koskinen (1992).

To this author's knowledge, the only other study to directly compare incidental vocabulary learning through watching videos and reading was conducted by Vidal (2011). 230 Spanish undergraduate EFL students were divided into three groups. Over the course of three weeks, one group watched three videotaped academic lectures (14-15 minutes each), one read texts on the same topics, and a control group received no treatment. The lectures and readings had the same content and target words; only the style was slightly altered to adapt them to each medium. The target words were 36 lexical items from the University Word list, which were repeated in each lecture/text between one and six times. Gains were assessed in pre-, post- and delayed post-tests (after one month) using a modified version of the Vocabulary Knowledge Scale (Wesche & Paribakht, 1996), which assessed form recognition, L2-L1 translation, recall of meaning, and syntactic knowledge. The adapted scale had been tested and declared to be reliable by Vidal (2003).

The immediate post-test indicated that both the videotaped lectures and the readings resulted in significant vocabulary gains. Though smaller, the scores on the delayed post-test were still significantly larger than pre-test scores, indicating long-term learning gains. The gains of the reading group were greater than those of the video group in terms of all four aspects of knowledge tested: on average, the reading group learned 19% to 38% and the video group learned 7% to 28% of the previously unknown words (depending on which aspect of knowledge was measured). The reading group also showed greater retention on the delayed post-test. However, the differences between the two groups decreased as the students' proficiency increased, indicating that higher-proficiency learners may be able to benefit equally from engaging with both media. There were no significant changes in the control group's scores from pre- to post-test, indicating that the experimental groups' gains could be attributed fully to the treatment. Finally, the results showed that the number of times target words occur in the input predicted 47% and 24% of the variations in the gains of the reading and listening groups respectively.

The generalisability of Vidal's (2011) results to incidental learning from other types of video is questionable. Firstly, academic lectures are not representative of the types of materials L2 learners might choose to engage with in their spare time or of the resources typically used in L2 classrooms. Secondly, it

is unclear what the visual aspect of the videos used in this study was, e.g. whether the videos showed the lecturer or whether there were slides or pictures. In any case, the author does not seem to attribute great effects to these visuals, as she refers to learning from the lectures as learning from *listening*. She does not discuss how watching a videotaped lecture differs from listening to an audiotape.

2.5. Variables That Affect Incidental Vocabulary Learning

In the review above, several variables have emerged that influence incidental vocabulary gains from reading and watching videos, such as the learners' proficiency (Hui, 2007; Neuman & Koskinen, 1992; Vidal, 2011; Zahar et al., 2001) and vocabulary size (Horst et al., 1998; Rodgers, 2013), the target words' general frequency (Horst et al., 1998), word class (Pigada & Schmitt, 2006), and in-text frequency (Horst et al., 1998; Pigada & Schmitt, 2006; Rodgers, 2013; Vidal, 2011; Zahar et al., 2001), as well as the contextual informativeness (Neuman & Koskinen, 1992; Vidal, 2011; Zahar et al., 2001). Two of these factors (vocabulary size and the target words' in-text frequency), which have been found to be especially influential, were controlled for in the present study (see section 4.4.). Therefore, previous research on their influence is discussed in this section in more detail.

2.5.1. Coverage and Vocabulary Size

Incidental vocabulary acquisition requires learners to have good comprehension of the input materials, so that they can use contextual clues to infer the meaning of unknown words (e.g. Hu & Nation, 2000; Liu & Nation, 1985). To reach a minimum level of comprehension, in turn, it is generally assumed that readers need a certain *lexical coverage* (ratio of known and unknown words in the text).

Research findings converge to indicate that a lexical coverage of about 95% to 98% provides learners with sufficient comprehension (e.g. Hirsh & Nation, 1992; Hu & Nation, 2000; Laufer & Ravenhorst-Kalovski, 2010; Laufer, 1989; Liu & Nation, 1985; Stæhr, 2008) and this is attained by a vocabulary of about 5,000 word families (Adolphs & Schmitt, 2004; Adolphs, 2003; Hu & Nation, 2000; Laufer, 1997; Nation & Waring, 1997; Nation, 1990, 2006). It must be noted that, even though these values are commonly called *thresholds*, they are more accurately interpreted as probabilistic boundaries, i.e. points at which it becomes

significantly more likely that a learner will understand the text (Hu & Nation, 2000; Laufer, 1989). Indeed, vocabulary size has been found to predict as much as 72% of the variance in reading comprehension (e.g. Henriksen, Albrechtsen, & Haastrup, 2004; Milton, Wade, & Hopkins, 2010; Stæhr, 2008). Moderate correlations have also been reported between vocabulary size and listening comprehension (e.g. Bonk, 2000; Milton et al., 2010; Stæhr, 2008).

The relationship between coverage and comprehension is often falsely equated with that between coverage and incidental vocabulary learning, although the latter relationship is less well explored. Horst, Cobb and Meara (1998, described in section 2.4.1.) reported that their participants' vocabulary size showed moderate correlations with their incidental learning gains (r = .31 and .36 respectively). Rodgers (2013), however, found that a coverage of above 94% lead to significantly better comprehension but did not influence incidental vocabulary gains. The influence of lexical coverage on vocabulary gains thus remains unclear to date. Yet, to control for a possible effect and to ensure adequate comprehension of the input material, the present study confirmed that all participants could reach approximately 96% lexical coverage (see section 4.4.3.).

2.5.2. In-text Frequency

The research discussed in section 2.4. indicated that, due to its gradual and incremental nature, incidental vocabulary learning requires repeated exposure to the same words. Widely different estimates have been proposed, however, for the exact number of encounters needed to produce significant learning gains.

Horst, Cobb, and Meara (1998; described in section 2.4.1.) found a significant correlation between the number of encounters with and learners' subsequent knowledge of a target word (r = .49). They suggested that participants should encounter words eight times or more to make 'sizeable gains' (Horst et al., 1998: 215). Similarly, Waring and Takaki (2003; described in section 2.4.1.) reported that learners needed to encounter a word at least eight times to have a 50% chance of recognizing its form or meaning three months later, and that having encountered a word eighteen times only gave learners a 10% to 15% chance of recalling its meaning after three months. For items occurring fewer than six times, this chance was almost zero. The authors extrapolated that 20 to 30 encounters could be needed to fully acquire a new word in the long term.

Webb (2007) designed a study specifically to test the effects of in-text frequency on incidental learning. He exposed 121 adult Japanese EFL learners to ten target words occurring one, three, seven or ten times in a reading task. The target words were replaced with pseudo-words to avoid pre-testing and vocabulary gains were measured using a battery of ten tests measuring recognition and recall of various aspects of vocabulary knowledge (orthography, syntax, meaning and form, associations, and grammatical functions). The number of encounters was significantly correlated with the size of vocabulary gains (r = .29 to .5 depending on the aspect of knowledge measured). While from one encounter, learners demonstrated some gains in recognition but not recall of all aspects; ten encounters were needed to make significant gains in both. It should be considered that in this study presented the target words in a series of unrelated contexts; in one coherent passage, fewer encounters might be needed.

In Zahar et al.'s study (2001; described in section 2.4.1.) the number of encounters also correlated significantly with item learning (r = .36). The best-learned words occurred in the text seven times, the least learned ones only two to three times. It is interesting to note that the target words' in-text frequency was more important for lower-proficiency learners (Groups 1-2), where frequency accounted for 15% of the variance, than at higher-proficiency levels (Groups 3-5) where it accounted only for 4% to 6%. However, as it is unclear how exactly Zahar et al. (2001) measured proficiency – they mention that it was through a general language test which included both production and comprehension tasks – their results cannot be directly compared the other studies discussed above.

Finally, Vidal (2011; described in section 2.4.3.) reported that the target words' in-text frequency predicted variations in the gains of their reading and video groups (47% and 24% respectively). This finding is important as it suggests that text-related factors may differentially affect learning from the two modalities.

Chapter 3: Research Questions and Hypotheses

The following research questions were addressed in the present study:

1. Engagement with which form of social media, video blogs or blog posts, leads to more incidental vocabulary learning?
2. Are there differences regarding which aspects of vocabulary knowledge are best learned from which input medium?
3. To what extent do learners use the unique features offered by the two social media platforms?
    a. How could this interaction relate to learning gains?
4. To what extent did the social media used in this study appeal to the participants? To what extent were they similar to those that users would choose to engage with independently?

Regarding Question 1, it is hypothesised that the participants will demonstrate greater total learning gains from watching vlogs than from reading blog posts. This prediction is based on the fact that video materials provide learners with a wealth of information in more than one modality (i.e. aural and visual input) whereas the blog posts provide mostly written input. Additionally, it is likely that the learners find the videos more engaging, leading to more concentrated processing and in turn to greater learning gains. Similar findings have been reported in a previous study by Neuman and Koskinen (1992).

With reference to Question 2, it is expected that there will be differences between the two media in terms of the types of vocabulary knowledge they promote. Specifically, since the videos used in this study did not include captions (although they could be activated as an optional feature), it would be expected that they would lead to fewer gains in orthographic knowledge than the blog posts. However, since the vlogs also provide engaging, multi-modal language input, it is predicted that they lead to greater gains in the other aspects of knowledge measured in this study, namely recall and recognition of meaning and grammatical function.

Question 3 and 4 are exploratory, wherefore no specific expectations were raised with regards to their answers.

Chapter 4: Methodology

4.1. Design

The main goal of the present study, which took a predominantly quantitative experimental approach, was to investigate English vocabulary acquisition from engaging with two types of social media: blogs and video blogs. Additionally, screen-capture videos were collected from a subsample of participants for a qualitative investigation of their interactions with the input media.

Most research on incidental vocabulary learning has employed pre- and post-tests to assess learning gains. There are, however, several issues with this methodology: Firstly, pre-tests may not be sensitive enough and therefore overlook that participants already possessed partial knowledge of the target words, or of some of their aspects. Secondly, encountering the target words during pre-testing may in itself contribute to their acquisition, thus falsifying the post-test results. Finally, pre-testing lexical knowledge draws participants' attention to the fact that the study is concerned with vocabulary acquisition, possibly leading to a more focused effort to learn unknown words during the treatment period.

An alternative way of ensuring that participants have no prior knowledge of the target words, which is adopted in the present study, is to replace them with newly invented pseudo-words. This strategy was previously employed, for example, by Waring and Takaki (2003) and Webb (2007, 2008). Since the participants have never encountered these words before, all vocabulary knowledge in the post-test can be attributed entirely to the input materials – assuming the pseudo-words are not cognates of words with similar meanings that the participants may know in other languages.

4.2. Participants

The participants were 84 volunteers, who completed the study online in response to a call for participants posted on social media pages associated with the *Nerdfighter* community (see section 1.2.). The participants consisted of 79 females and 5 males, which reflects an imbalance in virtual communities at large; for example, women make up about 72% of the Nerdfighters (Green, 2014). The participants were aged 14 to 25 years (M = 20.2, SD = 2.8) and spoke 19 different

first languages, with German (*n* = 23) and Dutch (*n* = 18) being the most common. None indicated that they were raised bilingually. All participants reported having learnt English for a minimum of 5 years prior to the study (*M* = 10.4, *SD* = 3.3).

4.3. Procedure

As Swanborn and De Glopper (1999: 263) noted, 'to assess [incidental learning] as validly as possible, studies should be performed under as natural a circumstance as possible'. Sockett and Toffoli (2012) similarly stress the importance of authenticity when studying Online Informal Learning of English. Therefore, in an attempt at making the experimental setting as similar as possible to regular engagements with blog posts and vlogs, all participants completed the study online using a web survey hosted by the digital research software Qualtrics ("Qualtrics Research Suite," 2014).

Firstly, the participants were informed that there were two versions of the study and asked to choose which one they would like to complete. The longer version required participants to download and use the free software Jing (2014) to record and share a screen-capture video of their engagement with the input materials. A pilot test with users drawn from the same population as the participants determined that the long version of the study would take approximately 45 to 60 minutes to complete. However, it was anticipated that such an amount of time would discourage many potential participants. They were therefore also given the option to complete the study without screen-captures. According to a pilot test with another four users, this short version would only take about 20 to 30 minutes to complete.

Which version of the survey the participants completed (long or short) was self-determined; within each of these options, however, they were assigned to two experimental conditions using Qualtrics' randomisation algorithm. The final numbers of participants in each of the four resulting groups are listed in Table 2 below. The small number of participants in the video group who also submitted screen-captures (*n* = 2) is likely a result of technical difficulties during data collection (see section 6.4.2.).

	Blog Group	*Video Group*	*Total*
No Screen-capture	31	36	67
Screen-capture	10	2	12
Total	41	38	79

Table 2: Final group sizes

All participants were first presented with extensive information about the purpose and procedure of the study and asked to give their consent for participation. So as not to bias them towards focussing on intentionally learning new lexical items where they might naturally not, the purpose of the investigation was not fully revealed at this stage. Instead, participants were informed that this study was about English learners interacting with and learning from English-language blogs and vlogs (see section 4.6.1, ethical considerations). To prevent them from discovering that the study used pseudo-words by looking up their meaning, they were instructed not to use other websites during the study.

Next, all participants completed a vocabulary test (section 4.4.2.). Those who had volunteered to record screen-captures were then provided with short instruction videos, in which the researcher modelled how the participants could show how they were interacting with the websites, e.g. by using the mouse cursor to follow along the lines of text or comments that they were reading.

After reading the blog posts or watching the vlogs, all participants completed a surprise vocabulary test that measured different aspects of their knowledge of six pseudo-words included in the videos and blog posts (section 4.5.). They also completed a questionnaire which asked for demographic background information and about their habits regarding blog posts or vlogs (Appendix 3). Finally, participants were debriefed about the study's full purpose and made aware of the fact that the words they had been tested on were pseudo-words.

4.3. Input Materials

In the first condition (henceforth *video group)*, the participants watched three vlogs, two to five minutes in length. In each video, one video blogger (vlogger) talked directly to the viewers about his or her experiences with an overarching topic: online friendships. Some photographs, which the vloggers referred to in their narration, were also shown. In the second condition (henceforth *blog*

group), participants read three blog posts that were created from the videos' scripts. Each posts was between 450 and 800 words in length (1,691 total) and contained the same photographs as were shown in the videos because without them, the text could not have been understood.

The blog posts and videos were created in collaboration with three vloggers, who volunteered to take part in this project. They can be classified as established YouTube creators, as prior to this study, they had been creating YouTube content for three to five years, had had 850, 1,500, and 15,5000 subscribers respectively and their videos had in turn been viewed a total of 80,000; 110,000; and 1,105,000 times[5].

Together, the vloggers chose the topic (online friendships) for their videos. They felt that they could easily talk about their own experiences regarding this theme and that it would appeal to a wide variety of viewers. Together with the researcher, the vloggers also chose the target words that would later be replaced by pseudo-words (section 4.4.). Next, each vlogger independently wrote a script for a video no longer than five minutes, as this is the recording limit allowed by the freeware Jing that participants would be using to record screen-captures. The scripts were then edited by the researcher to ensure consistent usage and a relatively even spread of the target words. Finally, the vloggers filmed the videos according to these revised scripts. The scripts also formed the basis for the blog posts (Appendix 1), with only slight alterations (e.g. 'write to you' instead of 'talk to you') being made to adapt to the different media format.

The videos and blog posts were then uploaded by the researcher to a YouTube channel ("OxEd," 2014) and Blogspot page ("International Friendships," 2014) created specifically for this study. Both were *private*, meaning they could only be accessed via hyperlinks provided in the web survey. These links were also distributed to a group of conspirators, who before the launch of the survey were asked to comment on the videos and blog posts, occasionally using the pseudo-words. This was done to create a more authentic online environment and to give starting points for the participants to engage in discussions in the comments sections.

[5] These data were retrieved from the three video bloggers' YouTube channel information pages in February 2014. The vloggers asked for their pages not to be cited.

4.4. Target Words

Ten target words, as in Webb's (2007) study, were initially selected but later reduced to six because the researcher considered that this would help to reduce the testing burden. Three nouns, two verbs, and one adjective were selected because these are the word classes most commonly found in natural speech, and because these numbers reflect their proportional frequency of occurrence in general speech (Kucera & Francis, 1967). The specific words were chosen because they were easy to repeat frequently when discussing international and online friendships but would not be vital to understanding the overall message. The six target words were *time zone, message, photo, (to) travel, (to) write,* and *interesting*.

4.4.1. Word Forms, Repetitions, and Context

The vloggers were instructed to always use grammatically ambiguous target words as the same part of speech; for example, although *travel* can be a noun or a verb, it was always to be used as a verb. While this was obeyed in the scripts, one vlogger accidentally once used *message* as a verb in his video. The mistake was not detected until after the editing process was concluded and time restrictions of the project prevented re-filming and -editing. This mistake was therefore carried over into the matching blog post and taken into consideration during scoring (see sections 4.5.3. and 4.5.4.).

In accordance with previous research (section 2.5.2.), it was determined that participants would need to encounter each target word at least ten times to achieve substantial learning effects. The vloggers were therefore instructed to incorporate each target word in their script a minimum of three and a maximum of six times (to avoid over-loading the scripts with the substitute words, thus making them overly salient). Each item was repeated 11 to 14 times in total (see Table 3). Participants were also able to encounter the target words more often, for example if they chose to re-watch or re-read passages or to read and write a comment on a text or video. Although previous research (section 2.5.2.) found that more frequent exposure correlates with better learning of target words, it was decided that the effect of the potential additional exposures would not be accounted for due to practical constraints imposed by the method of data collection. Similarly, even though contextual informativeness has been found to

correlate to learning gains (Neuman & Koskinen, 1992; Vidal, 2011; Zahar et al., 2001), the present study did not manipulate the contexts in which the target words were presented, so as not to compromise the authenticity of the vloggers' own style.

4.4.2. Pseudo-Words

The pseudo-words' orthographic forms were adopted from Webb (2007), who, in order to retain face validity as possible existing words in English, created them to conform to normal English conventions in spelling and phonology. Webb (2007: 52) also reported that in a pilot study, intermediate ESL learners believed the pseudo-words to be authentic English words. The target words and substitutes are listed in Table 3. In the blogs and videos, the words were given standard English inflections where needed.

Word Forms		*Number of Repetitions*[6]			
Target Word	*Pseudo-Word*	*Script 1*	*Script 2*	*Script 3*	*Total*
time zone	*pacon*	3	4	4	11
message	*masco*	4	4	5	13
photo	*sagod*	4	4	6	14
(to) travel	*tasper*	4	4	5	13
(to) write	*ancon*	4	4	5	13
interesting	*dangy*	4	4	4	12
Total		24	23	29	76

Table 3: Repetitions of pseudo-words

4.4.3. Coverage and Comprehension

It is generally assumed that learners must be able to understand a certain proportion of words in the input to reach sufficient comprehension and be able to learn new vocabulary from it. Research findings converge to indicate that a lexical coverage of about 95 to 98% provides learners with sufficient comprehension and that this is attained by a vocabulary of about 5,000 families (see section 2.5.1.). To estimate whether participants would be able to reach sufficient coverage in this study, their general lexical knowledge was assessed using the

[6] Full scripts provided in Appendix 7.

5,000-word Vocabulary Levels Test Version 1 (VLT; Schmitt, 2000; see also Schmitt, Schmitt, & Clapham, 2001 for validity testing).

Five participants who answered fewer than 23 out of 27 items (85%) correctly were excluded from the analysis. This criterion was chosen to approximate Schmitt et al.'s (2001) criterion of mastery of a vocabulary level (26 out of 30; 86.7%). The scores of the remaining 79 participants (M = 25.8, SD = 2.2) indicated to the researcher's satisfaction that they had receptive knowledge of a majority of the 5,000 most frequent English words and therefore should have no problem inferring the meanings of the novel target words.

To support this assumption, a frequency profile of the input materials was created using the Lextutor Vocabprofile software (Cobb, 2013). When excluding proper nouns from the analysis, but including the pseudo-words, only 13 out of 337 lemmas in the scripts (3.9%) were not part of the 5,000 most frequent words in English. The input materials could thus be assumed to be comprehensible to the participants, as they would have at least 96.1% lexical coverage.

4.5. Dependent Measures and Scoring

Immediately after engaging with the input materials, the participants were given a range of tests designed to measure different aspects of vocabulary gains (Appendix 2). Delayed post-tests were not administered due to the time-constraints of the study and because, as Webb (2008: 237) remarks, 'without the possibility of further encounters with the [pseudo-words ...], the rate of decay of vocabulary knowledge should follow established norms'. (See for example Waring & Takaki (2003; described in section 2.4.1.), who investigated the decay of pseudo-word gains).

Each test was presented on a new page of the web survey and participants could take as much time as they needed to finish. The measures were adapted from Webb (2007, described in section 2.5.2.), who designed a test battery to examine the recognition and recall of five aspects of vocabulary knowledge: orthographic form, meaning, grammatical function, syntactic function, and paradigmatic associations (e.g. semantic category membership, synonymy, antonymy, etc.). It was expected that the participants would perform better on the recognition tests as these were more sensitive to partial knowledge: they were multiple-choice tests in which the distractors were chosen to be very

different from the correct answer so that the participants would be able to choose the correct answer even if they possessed only partial knowledge of the target word. The recall measures were designed to be more difficult and require fuller knowledge of the target words.

To reduce the testing burden, only five out of Webb's ten tests were included in the present study. The measures of syntactic knowledge were dropped because they measured an aspect of knowledge that seemed sufficiently covered by the tests of grammatical function (sections 4.5.3. and 4.5.4.). The tests of paradigmatic associations were excluded because Webb (2007) noted in his discussion how the instructions for this test, even when explained in the participants' L1, were frequently misunderstood, making it an invalid measure. Finally, the test of recognition of orthographic form was also omitted. By allowing for small deviations while scoring the test of orthographic recall, the researcher felt that participants were given sufficient opportunity to demonstrate partial knowledge (see section 4.5.1.).

The remaining tests were presented in the same order as in Webb's (2007) original study. They were 'carefully sequenced to avoid earlier tests affecting answers to later tests' (Webb, 2007: 54). The testing methodology was not altered with the exception of the tests of knowledge of meaning (sections 4.5.4. and 4.5.4.). In Webb (2007: 56–58), these were L2-L1 translation tests. In the present study, however, not all participants spoke the same L1, meaning that the researcher would not have been able to score such a test reliably. Therefore, monolingual tests of meaning recognition and recall were used, the former adapted from Webb's (2008) study.

All items were presented without contexts, which may have cued recall and provided clues (e.g. about the words' grammatical function) for the upcoming tests. To help suspend participants' possible disbeliefs about the realness of the pseudo-words (see discussion section 6.3.), all tests also included four distractors. (It was felt that more would have made the testing procedure too time-consuming.) These distractors were low-frequency words appearing only once in the input materials*: acquire*, *unfortunate*, *poutine*, and *collab* (net-speak for collaboration). The answers for these items were not analysed. Finally, to control for a guessing-bias (cf. Schmitt, 2010), participants were told in each test that

they should try to answer if they thought they might know the words, but not to guess if they did not and instead answer 'I don't know'.

The five dependent tests are described below in the order in which they were presented to the participants.

4.5.1. Orthography

The first test (Appendix 2.1.) measured participants' productive knowledge of the pseudo-words' orthographic forms. Participants were presented with one audio recording for each item, in which a male British English native speaker read out the word twice. They were asked to write each word in a text field below the audio player. Their answers were marked as correct if they contained no spelling errors. However, other plausible forms with minor spelling alternations (e.g. 'paycon' for *pacon;* 'sagot' for *sagod*, etc.), were also accepted to allow for partial knowledge, but not if they spelled other English words (e.g. 'bacon', 'icon', 'pecan', etc.). Two English native speakers (one British and one North American) judged whether the variations were plausible by English orthographic standards.

4.5.2. Recall of Meaning

This test (Appendix 2.2.) asked participants to write a short description of the meaning of each pseudo-word and distractor. Two examples were provided to illustrate the kinds of answers that were sought. Answers were marked as correct when they constituted the word that had been substituted by the pseudo-words (e.g. 'time-zone' for *pacon*) or a circumlocution of its meaning (e.g. 'written communication' for *message*). Since the goal of using a variety of measures was to record partial knowledge, paraphrases and synonyms were also accepted, even if they matched only one of the target words' senses, e.g. 'cool' for *interesting* or 'compose' instead of *write*.

4.5.3. Recall of Grammatical Function

The third test (Appendix 2.3.) instructed participants to write sentences containing the target items. Again, the kinds of answers that were sought were illustrated by two examples. An answer was marked correct if the target item was used in context as the correct part of speech (noun, verb, or adjective). Whether

or not the sentence reflected an understanding of the word's meaning was not taken into consideration at this stage. For example, for the pseudo-word *pacon*, both 'My boyfriend and I live in different pacons' and 'I gave him my pacon yesterday' would have been marked as correct, even though the latter context is not appropriate for the meaning of *pacon* (time-zone). However, 'He pacons every day' would have been marked as incorrect because it uses *pacon* as a verb instead of a noun.

Special rules had to be applied in scoring the answers for one of the items, *masco*. As has been mentioned before, while this pseudo-word was supposed to replace the noun *message*, the blogs and videos contained one instance of it being used instead of the verb (*to) message*. Any answer in which *masco* was used as a verb was therefore also marked correct. Therein, it was assumed that identifying *masco* as a verb was indeed due to its usage in that function the input and not because the corresponding English word (*message*) can be both a noun and a verb; in the latter case, *tasper* (to travel) should in all likelihood also have been often misidentified as a noun (travel), which did not happen.

4.5.4. Recognition of Grammatical Function

In this multiple-choice test (Appendix 2.4.), each item was presented alongside three simple sentences, in which it was used in a different grammatical function: once as a noun, once as a verb, and once as an adjective. Participants were asked to identify the sentence in which the item was used correctly. As with the test of recall of grammatical function, *masco* was marked correct for being identified either as a noun or a verb.

4.5.5. Recognition of Meaning

In this final multiple-choice test (Appendix 2.5.) participants had to identify, for each item, which option best described its meaning. The options included the original target words (which had been replaced with pseudo-words) as well as three distractors items. The distractors were of the same parts of speech as the target words and were plausibly related to the topic discussed in the input materials (online friendships). They were also selected to have roughly the same frequency in general speech as the target words. These frequencies were determined using the Corpus of Global Web-Based English (Davies, 2014), composed of 1.9 billion words from 1.8 million English-language websites. This

corpus was chosen because it provides frequencies of use in web-based text (a context similar to that of the present study). However, because this corpus did not allow searching for words (potential distractors) by frequencies, its data was triangulated with the SUBTLEX-US (Brysbaert & New, 2009), which includes frequencies of over 75,000 words based on the subtitles from American films and television series. This corpus was chosen because, like the materials in this study, it is concerned with language used in video materials.

4.6. Ethical Considerations

This project followed the ethical guidelines proposed by the British Educational Research Association ("Ethical Guidelines for Educational Research," 2011). Prior to data collection, approval was obtained from the University of Oxford Inter-divisional Research Ethics Committee by completing the CUREC/1A Checklist for the Social Sciences and Humanities and the CUREC/2 Form. The main concerns with this research, which warranted the submission of the CUREC/2, were issues of participant information and consent, online safety, and confidentiality and anonymity. These will be discussed in turn below.

4.6.1. Participant Information and Consent

So as not to bias participants towards intentionally focusing on vocabulary learning where they might naturally not, it was essential that the participants were not aware that they would be tested on their acquisition of words included in the blogs and videos. The participants were therefore deceived about the true purpose of the study in the digital participant information sheet. They were retrospectively debriefed about the study's full purpose immediately after completing all tests and questionnaires.

In order to be representative of the population from which the participants were drawn (the Nerdfighter community), the study recruited participants between the ages of 14 and 25. Participants under the age of 18, however, are seen as unable to give voluntary informed consent ("Ethical Guidelines for Educational Research," 2011: 6). Ordinarily, consent would be obtained in these cases from the underage participants' parents or legal guardians. In the present study, however, such practice was not feasible because the research was conducted online with participants from a variety of countries.

All participants were therefore asked to complete a digital consent form and to indicate their age and date of birth. Participants reporting that they were less than 18 years old were asked to confirm that their parents/guardians knew about their participation in the study and were happy for them to take part. It was assumed that participants would be honest concerning their age and concerning parental consent, although it could not be ruled out that individuals older or younger than the target range would choose to participate in the study, or that underage participants might choose to do so without parental consent. Should this have occurred, however, there would have been no particular risk to the individuals, as all content presented in the study was suitable for all ages and precautions were taken to ensure the participants' online safety as detailed below. It was stressed that participation was voluntary and that they would be able to withdraw from the study at any time and their data subsequently deleted. Participants were also invited repeatedly to contact the researcher or supervisor if they had any questions or concerns.

4.6.2. Online Safety

Because participation in the study involved engaging with other users online, some added measures were taken to guarantee the online safety of all participants, especially those under the age of 18.

Both social media platforms used in this study, Blogspot and YouTube, are owned by Google and display users' real names (from their profiles on the social network Google+) instead of usernames by default. When commenting on the videos or blog posts, participants were therefore encouraged to do so as anonymous guest users or by manually changing their display name to a user name of their choice. Additionally, all participants were reminded of the importance of safe internet conduct and provided with a link to an online resource about safe interaction on social networking sites ("Social Networking Sites," 2012).

To protect both the participants and the vloggers and to prevent unwanted contact between them, the vloggers were identified only by their first names (no last names or YouTube usernames). Regardless, the vloggers could be assumed to be 'safe' for young participants should contact still be made, since their large and

established following made it likely that safeguarding issues would have come to light already, had there been any.

The videos and blog posts were uploaded to a YouTube channel ("OxEd," 2014) and Blogspot page ("International Friendships," 2014) created specifically for this study. Both were *private* pages, meaning that they could only be accessed via secret hyperlinks but were not open to the public. These links were accessible only to the researcher, the participants who consented to take part in the study, and a group of conspirators, who commented on the YouTube videos and blog posts before the launch of the survey.

Finally, the participants were repeatedly encouraged to contact the researcher or supervisor in the unlikely event that they encountered any problems or had any concerns.

4.6.3. Confidentiality and Anonymity

The web survey used to collect data ("Qualtrics Research Suite," 2014) saved each participant's data in a separate PDF file. These included not only the participants' answers to the VLT, questionnaire, and dependent measures but also their IP addresses (which Qualtrics automatically records) and email addresses, if the participants chose to supply them for the prize draw or to hear about the results of the study. The researcher subsequently removed all identifying information from the files. Every file was assigned a participant number, which was used to identify participants throughout the study. The email addresses were extracted from the answer sheets and stored in a separate file.

The screen-capture videos, which some participants submitted to the researcher, however, could not be anonymised. If a participant chose to comment on a video or blog post, their username would be visible in the video. This was unavoidable due to the method of data collection, but this information was not used or disclosed in any part of the research.

The researcher kept all records securely and deleted the data from Qualtrics upon completion of the study. All comments (by conspirators and participants) were also deleted from the YouTube and Blogspot sites at the conclusion of the study. The participants themselves were responsible for deleting the screen-capture videos from their accounts on Screencast, the digital

database connected with the Jing screen-capturing software. They were reminded to do so in the participant information and consent documents.

Chapter 5: Results

To determine whether parametric tests were appropriate for the quantitative analysis, all interval variables were tested for the Normality of their distribution. This was determined by a visual inspection of the histograms and by calculating values of skewness and kurtosis: According to conventions in Social Science research (Cameron, 2004), a variable was considered Normal if both skewness and kurtosis fell within the range of -2 to +2.

The significance level for the inferential tests reported below was initially set at $\alpha = .05$. However, a Bonferroni adjustment was applied to correct for the effect of multiple comparisons. A total of ten comparisons of means were conducted; hence the critical level of significance was set at $\alpha = .005$. When data were missing, cases were excluded pairwise and relevant sample sizes are reported. Pairwise exclusion was chosen to maximise statistical power across all analyses, given the reductions of sample sizes that would have resulted from listwise exclusion.

5.1. Comparing the Experimental Groups

Before the main analysis, the two experimental groups were compared in terms of demographic variables (age, L2 experience in years, and level of education) and Vocabulary Levels Test (VLT) scores to ensure that they were sufficiently similar. The differences between the groups in terms of the Normal interval variables were calculated using a series of independent t-tests. Levene's Test for the equality of variances was non-significant in all cases. For the ordinal Level of Education variable, a Mann-Whitney U test was employed. The results are reported in Table 4. There were no significant differences between the groups in terms of any of the variables tested.

Variable	*Range*	*Distribution*				
		Video	*Blog*	*t*	*df*	*p*
Age	14-29 years	*M* = 19.95, *SD* = 2.27 *n* = 38	*M* = 20.51, *SD* = 3.28 *n* = 41	-.88	77	.38
VLT Total	23-27 out of 30	*M* = 26.24, *SD* = .97 *n* = 38	*M* = 26.17, *SD* = 1.26 *n* = 41	.26	77	.80
L2 Experience	2-20 years	*M* = 10.19, *SD* = 3.25 *n* = 37	*M* = 10.58, *SD* = 3.45 *n* = 40	-.51	75	.62
Variable	*Range*	*Video*	*Blog*	*U*		*p*
Level of Education	0 (less than High School) to 4 (Master's degree)	*Mdn* = 2 *n* = 37	*Mdn* = 2 *n* = 41	637.5		.20

Table 4: Distribution of demographic variables and VLT scores

5.2. Research Question 1: Total Gains

The first research question asked engagement with which of the two types of social media, if any, would lead to greater incidental vocabulary gains. For each participant, the total vocabulary gains were calculated by adding the scores of all six items (0 = false, 1 = correct) from the five sub-tests of vocabulary knowledge. This new interval variable was thus measured on a scale from 0 to 30 and its distribution (see Figures 1 and 2) was found to be Normal in both groups.

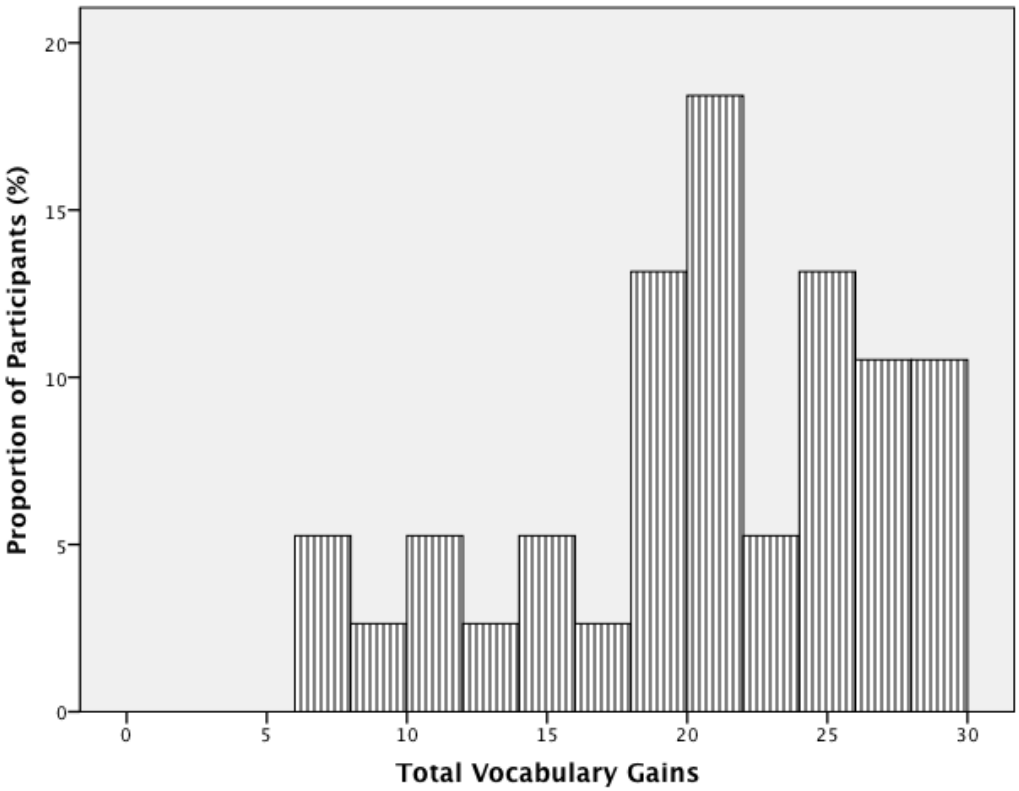

Figure 1: Distribution of total gains in the video group

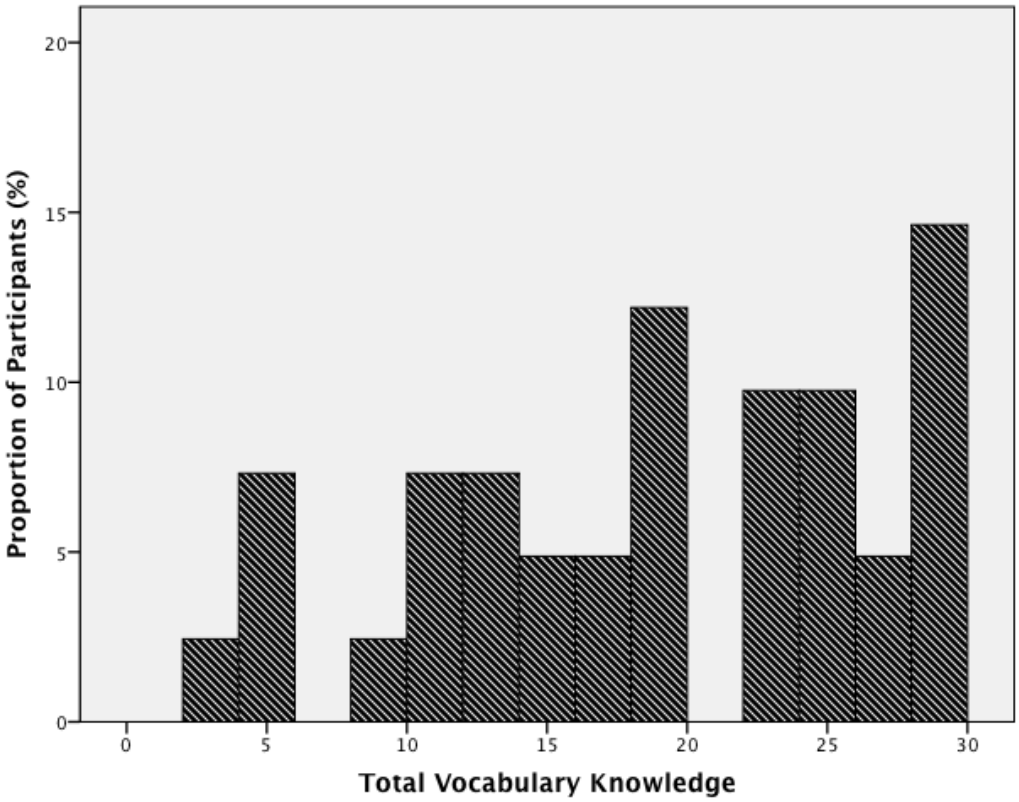

Figure 2: Distribution of total gains in the blog group

An independent samples t-test indicated that there was no significant difference between the video group (M = 20.74, SD = 6.35, n = 38) and the blog group (M = 19.76, SD = 8.37, n = 41) with regard to their total vocabulary gains, $t(73) = .58$, $p = .56$, $d = .13$. However, Levene's Test indicated unequal variances ($F = 5.75$, $p < .05$), so the degrees of freedom were adjusted from 77 to 74. As illustrated by Figure 3 below, if we consider the maximum score on the overall gains measure (30 points), we could say that average gains ranged from 58% (25th percentile) to 87% (75th percentile) of previously unknown words for the video group and from 42% (25th percentile) to 93% (75th percentile) for the blog group.

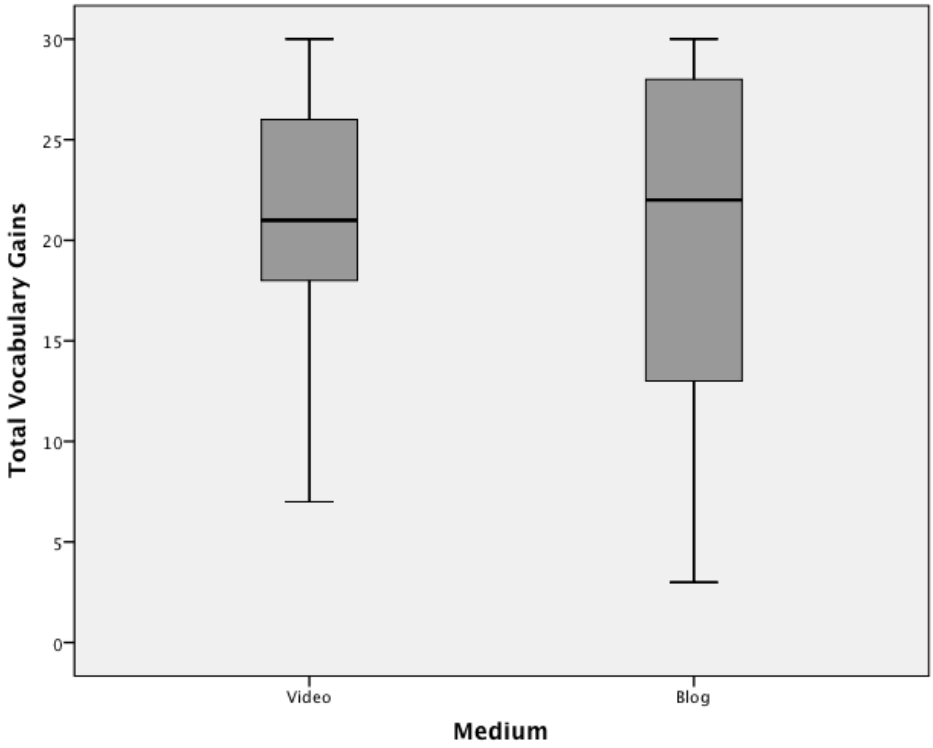

Figure 3: Distribution of total gains from the two media

5.3. Research Question 2: Aspects of Vocabulary Gains

The second research question asked whether there were differences regarding which aspects of vocabulary knowledge are best learned from watching vlogs and reading blog posts.

Appendix 4 contains histograms representing the distribution of gains in both groups regarding the five aspects of vocabulary knowledge measured. The two groups' mean scores (out of six) on these measures were compared using a

series of independent samples t-tests. The results are reported in Table 5. There was one statistically significant difference between the two groups: the blog group showed greater gains than the video group in terms of orthographic knowledge. Cohen's *d* indicated that this was a medium-size effect (*d* = .72).

Cohen's *d* furthermore indicated a medium-size effect of the input medium on the test of meaning recognition (*d* = .43) and two small effects on the recall of grammatical function and meaning (both *d* = .25), although these differences did not reach statistical significance. On all of these measures, the video group scored higher than the blog group.

Variable	*Distribution*					
	Video	*Blog*	*t*	*df*	*p*	*d*
Orthography	*M* = 3.71 *SD* = 1.66 *n* = 38	*M* = 4.79 *SD* = 1.30 *n* = 38	-3.16	74	.002*	.72
Grammatical Function (Recognition)	*M* = 4.26 *SD* = 1.41 *n* = 38	*M* = 4.07 *SD* = 2.07 *n* = 41	.48	71[a]	.63	.11
Grammatical Function (Recall)	*M* = 4.11 *SD* = 1.69 *n* = 36	*M* = 3.61 *SD* = 2.20 *n* = 41	1.13	73[b]	.26	.25
Meaning (Recognition)	*M* = 5.13 *SD* = .78 *n* = 38	*M* = 4.49 *SD* = 1.98 *n* = 41	1.93	53[c]	.059	.43
Meaning (Recall)	*M* = 3.74 *SD* = 1.86 *n* = 38	*M* = 3.23 *SD* = 2.24 *n* = 40	1.10	76	.28	.25

[a] Levene's Test indicated unequal variances ($F = 8.57$, $p < .05$), so df was adjusted from 77 to 71
[b] Levene's Test indicated unequal variances ($F = 5.72$, $p < .05$), so df was adjusted from 77 to 73
[c] Levene's Test indicated unequal variances, ($F = 17.42$. $p < .05$), so df was adjusted from 77 to 53
* significant at the corrected level $p < .005$

Table 5: Gains of five aspects of vocabulary knowledge

5.4. Research Question 3: Interactions with the Social Media

The third research question was concerned with exploring to what extent learners make use of the unique features offered by the two social media platforms in this study, Blogspot and YouTube.

A look at the comments sections of the blog posts revealed that five participants wrote a total number of seven responses, two or three on each post. The pseudo-words were used by participants in their own comments eight times, each an instance of correct use. On the videos, four participants left a total of five comments, with one vlog receiving three and the others receiving one response each. None of these comments included the pseudo-words. Content-wise, in all comments participants recounted their own experience with online friendships, sometimes in response to specific issues addressed in the videos or blogs (e.g. 'keeping in contact is so hard for me [...]. I'm glad to know I'm not the only one').

Most participants who submitted screen-captures (eight out of ten in the blog group, two out of two in the video group) scrolled through the comments section of all videos or blogs, indicating that they likely read or at least skimmed other users' comments. It could not be discerned, however, how many comments each participant actually read.

The blog group's screen-captures also revealed that participants scrolled through, and thus presumably read, the blog posts at widely varying paces. There were no instances of readers noticeably skipping (scrolling) back and forth in the text (which could have indicated pre- or re-reading of certain sections), except in cases where it was clear that the scrolling was to look at the photographs referenced in the text. The two participants in the video group, who recorded screen-captures, also did not skip back and forth in the video, pause, or re-wind it. Neither did they choose to turn on the English captions, which were provided optionally with each video.

5.5. Research Question 4: Appeal and Familiarity of Input Materials

Finally, the fourth research question asked about the extent to which the materials used in this study were similar to those that users would choose to engage with independently. The participant questionnaire (Appendix 3) asked three questions to explore this: whether the participants liked the videos and blog posts, whether they found the topic of online friendships interesting, and how

similar the format of the materials was to those that they usually engaged with. All were answered on a five-point Likert scale ('not at all' to 'very').

Figure 4 summarises how much participants in both groups reported liking the input materials. This variable follows a roughly similar, Normal distribution in both groups.

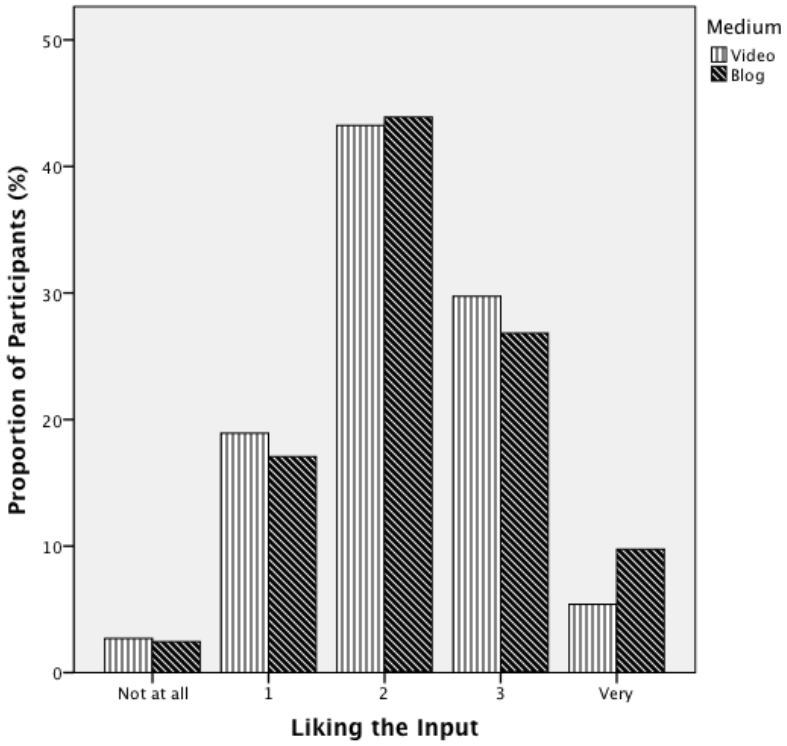

Figure 4: Did you like the videos you just watched/posts you just read?

Figure 5 shows how interesting participants in both groups found the topic of the blogs and videos. While a large proportion of participants in the video group (47%) found the topic neither interesting nor uninteresting (two points on the five-point scale), participants in the blog group seemed more inclined to report moderate levels of interest (34% choose two and another 34% choose three on the five-point scale).

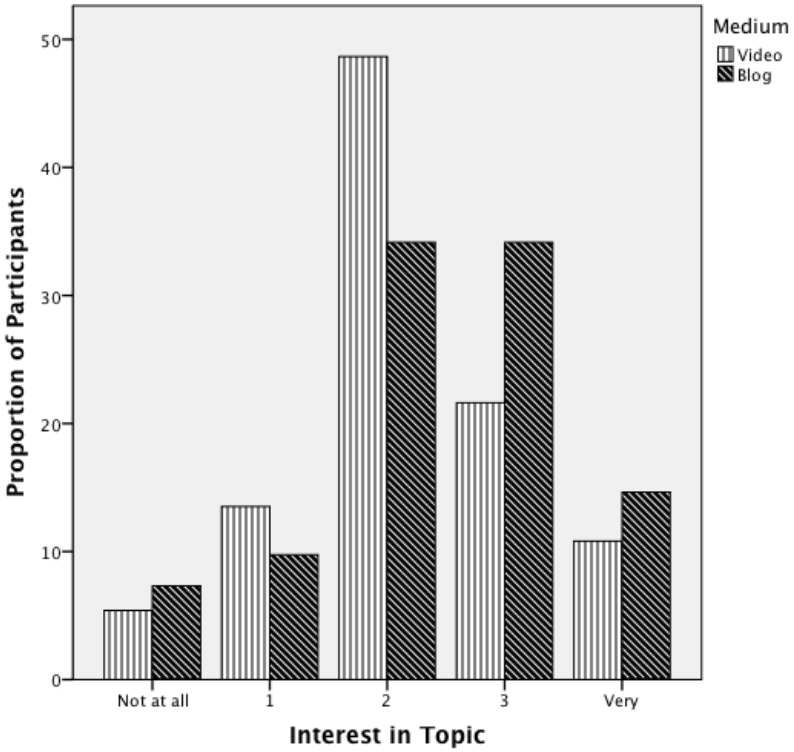

Figure 5: How interesting did you find the topic of the videos/blog posts?

Figure 6 indicates that the same content presented in the form of blog posts was less familiar to the participants than as videos.

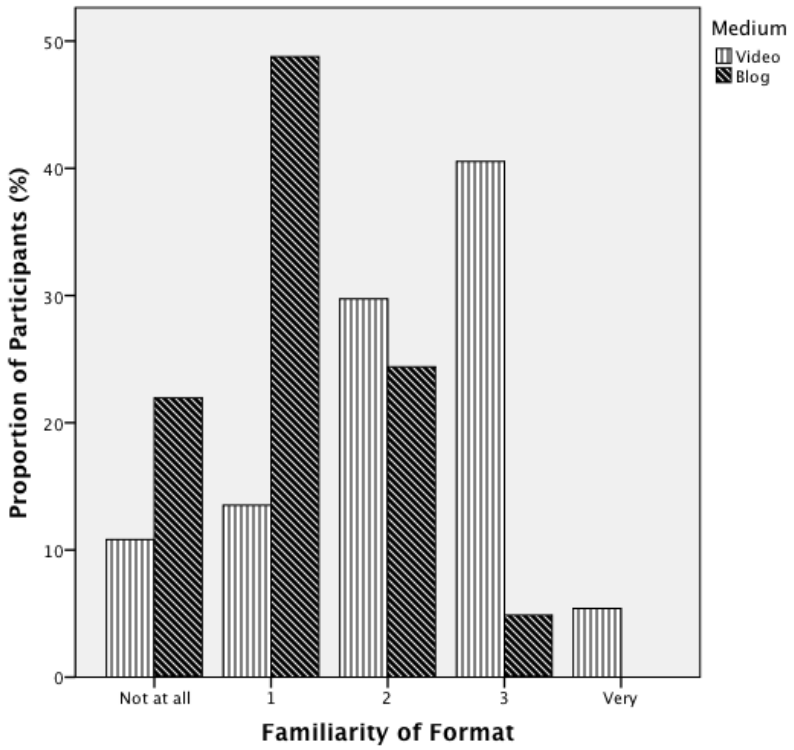

Figure 6: To what extent where the video blogs/blog posts like those that you usually watch/read?

Finally, the researcher was also interested in the extent to which the participants habitually engaged with this type of social media. Both groups reported frequent engagement (Figure 7). 60% of participants in the video group reported watching vlogs daily and 30% watched them two to three times per week. In the blog group, 81% claimed to read blogs every day and 12% reported reading them two to three times per week.

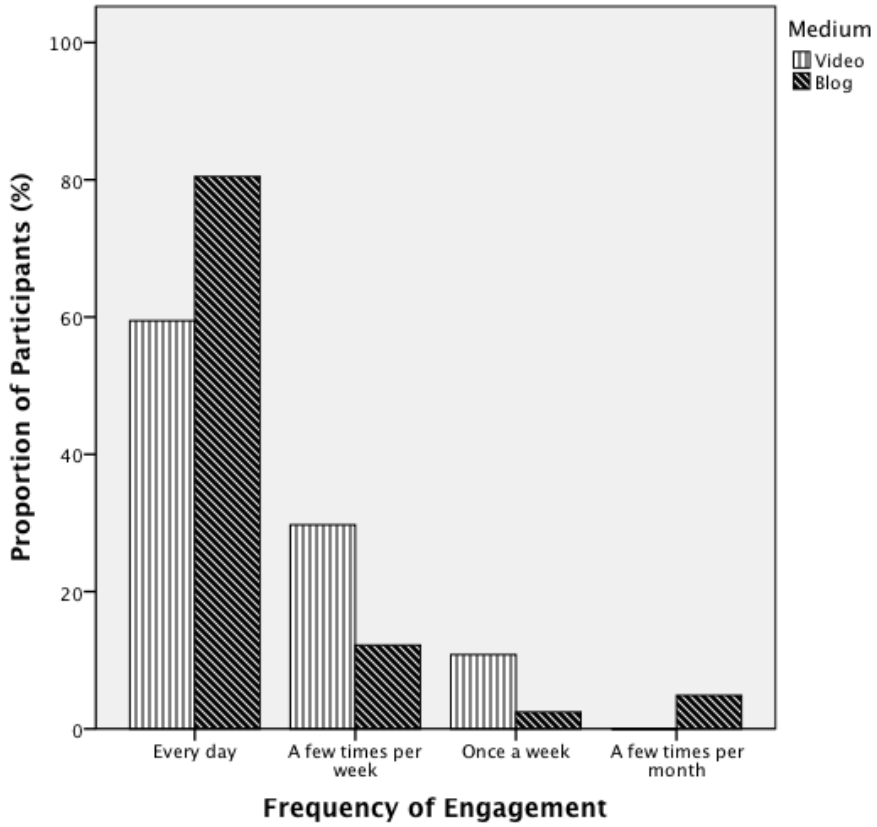

Figure 7: How often do you read blogs/watch vlogs?

Another figure that can give an idea of engagement with these media is the number of blogs or YouTube channels users follow (Figure 8). Most participants reported subscribing to between one and ten blogs or YouTube channels (61% and 71% respectively). Only few indicated that they subscribed to more than fifty, although this was more frequent in the blog than video group (20% versus 5%).

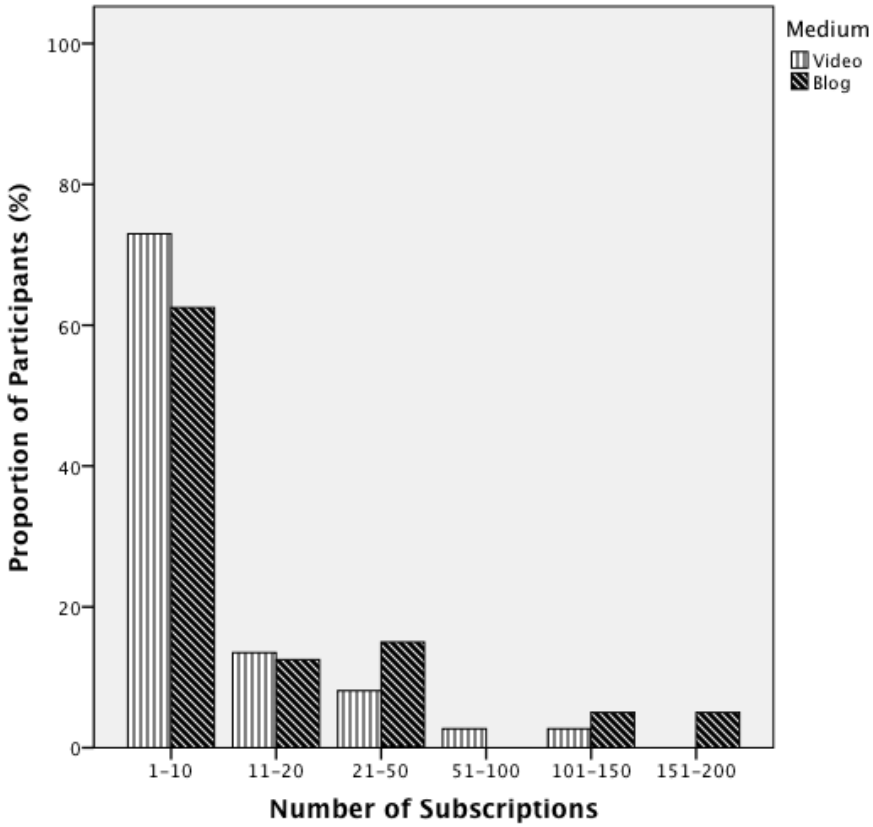

Figure 8: How many channels/blogs do you subscribe to?

Chapter 6: Discussion

6.1. Research Question 1: Total Gains

The total vocabulary gain scores support the findings in earlier studies that both text and videos can serve as suitable input for incidental vocabulary learning (among many others: Horst et al., 1998; Neuman & Koskinen, 1992; Pigada & Schmitt, 2006; Rodgers, 2013; Vidal, 2011; Zahar et al., 2001). It is difficult to make direct comparisons between the amounts of knowledge gained from these blogs and videos and other media in previous studies because of the differing lengths of treatments, numbers of items, and types of tests used. Therefore, one goal of this study was to directly compare the effects of the two types of media in a controlled experimental setting.

The mean overall vocabulary gains were slightly higher from watching the three vlogs (69% of target words) than from reading the blog posts (66%) but this difference was not statistically significant. However, it is interesting to note that there was significantly higher variation in the scores of the blog than the video group (see Figure 3, section 5.2.). Levene's Test of equal variances was violated, $F = 5.75$, $p < .05$.

One plausible explanation for this could be the way that these two types of materials were processed. The Depth of Processing Theory (Craik & Lockhart, 1972; Craik & Tulving, 1975) states that the extent to which a learner engages with the language input stands in direct relation to the learning she will be able to obtain. The blog posts could have led to a wider range of gains by allowing for more variations in the depth of processing: some readers may have only skimmed the texts and thus made fewer gains than those who read them in-depth. The screen-captures confirmed that participants in the blog group scrolled, and therefore presumably read, through the texts at widely varying paces, so that some could have processed the texts more deeply than others. Because it is not possible to skim a video for its main message in the same way, such wide variations in the depth of processing presumably would not have occurred in the vlog group, providing participants remained focused throughout the whole video. This explanation is supported by the fact that previous studies of incidental vocabulary gains from reading have let their participants listen to the text while

reading along to ensure equal levels of processing (e.g. Horst et al., 1998; Zahar et al., 2001).

Alternatively, the higher variability of learning gains within the blog group could be because other learner- and text-related variables, such as L2 proficiency or contextual informativeness, could differentially affect learning from the two types of media (cf. Vidal, 2011; section 2.5.2. who found that the target words' in-text frequency predicted more variation in the gains from reading than watching video).

6.2. Research Question 2: Aspects of Vocabulary Gains

The second research question addressed the possibility that the two types of input materials show differences in the particular aspects of vocabulary gains they promote.

Indeed, the input medium had a statistically significant, medium to large effect (d = .72) on the gains of orthographic knowledge. As expected, the blog group, who processed the target words in their written form, out-scored the video group on this aspect of knowledge. The input medium furthermore had a medium-sized effect (d = .43) on meaning recognition and small effects (both d = .25) on the recall of grammatical function and meaning. The presence of several effects with opposite directionality explains why the video and blog groups did not show significant differences in their total vocabulary gains.

The latter three effects were not statistically significant (all p between .059 and .28), which may be attributable to the fact that P-values, but not effect sizes, are sensitive to small sample sizes. The present study's relatively small sample (n = 79) resulted from the project's time restrictions as well as limitations placed by Qualtrics Research Suite (2014) on the number of participants who can answer to one survey (set to a maximum of 500). This had not been taken into consideration during the design of the study because Qualtrics only implemented this limit (as part of a series of changes to student accounts) after the data collection had already begun. Anyone who read the participant information or began to fill out the survey but dropped out was counted as a respondent. Therefore the data collection had to be stopped after the limit of 500 respondents was reached and amounted to only 84 full responses (instead of the target of 100), of which another five were excluded from analysis on the basis of their VLT scores. The

same study conducted with more participants may have held enough power to find these effects statistically significant.

The video group outperformed the blog group on all three aforementioned tests (recognition and recall of meaning, recall of grammatical function) tentatively supporting the hypothesis that watching videos can lead to better acquisition of some aspects of vocabulary knowledge than reading. Therein, the present analysis replicates Neuman and Koskinen's results (1992; described in section 2.4.3.), who found greater vocabulary gains from watching a science TV show than reading academic texts, and contradicts those of Vidal (2011; described in section 2.4.3.), whose participants showed greater lexical gains from reading academic texts than watching video-taped lectures. The contradictory findings could indicate that the vocabulary gains depend not only on the format, but also on the genre or topic of the input materials. This conclusion is also supported by Markham's (1999) study, in which participants made greater vocabulary gains from a video on American civil rights than from a documentary about whales.

6.3. Generalisability of the Recorded Gains

The total vocabulary gains reported in this study (*M* = 69% in the video and *M* = 66% in the blog group) were high compared to those reported in previous research (see section 2.4.). It is possible that the present study misjudged the vocabulary gains from reading blog posts or watching video blogs. Some reasons why this study's findings may be over- or underestimations of gains that can be expected from authentic Online Informal Learning of English are discussed below.

The use of pseudo-words ensured that the target items would be unknown to the participants prior to the treatment, therefore eliminating the need for pre-testing and ensuring that there were a certain number of new lexical items in the input for the participants to learn. It has previously been argued that acquisition studies using pseudo-words simulate authentic learning because they, like most L2 vocabulary acquisition, involve assigning a new label to a meaning that is already familiar (Service, 1992) and because 'form-meaning relationships are essentially arbitrary' anyway (Hulstijn, 2001: 262). As the participants in the

present study had not been informed that pseudo-words would be used, and Webb (2007) reports that ESL learners believed the same forms as those in the current study to be real English words, authentic learning was expected to occur.

However, some participants in the present study (approximately 10%) indicated that they knew that the pseudo-words were not real. During the meaning recall task, for example, one participant wrote 'I think you invented this word, but I think it means...'. It must therefore be considered what effect the pseudo-words may have had on learning gains.

Some researchers have argued that the use of pseudo-words may lead to an underestimation of learning gains: Papagno, Valentine, and Baddeley (1991), for example, suggested that learners who are aware of the use of pseudo-words would be less motivated to learn them because they would useless in the real world.

One reviewer of Webb's study (2008: 241) suggested that using pseudo-words for already known words may reduce the learners' chances of correctly inferring their meaning from context (because they already know another word that expresses the same meaning), leading to an underestimation of learning gains. Liu and Nation (1985), on the other hand, argued that, if participants already knew the original words, clues like collocational knowledge may help them guess the meaning of the pseudo-forms, thus leading to an overestimation of knowledge gains. Relatedly, it must be noted that in the present study, the pseudo-words replaced words with relatively high frequencies in general speech, which could have made it easier for the participants to infer their meaning.

Finally Pulido (2007) reasoned that recognising the use of pseudo-words goes along with identifying the gap in one's own knowledge about the meaning of these items, which according to the Noticing Hypothesis (Schmidt, 1990) is a condition for learning and could also lead to an overestimation of gains. The two contradicting predictions, an underestimation or overestimation of learning gains due to the use of pseudo-words, suggest that the effects of using pseudo-words in vocabulary learning research require further investigation.

Another feature of the input materials that could limit the generalisability of the reported learning gains was that each pseudo-word was embedded in each video or blog three to six times (eleven to fourteen times in total). This was done to control for the known effects of the in-text frequency of target words on

learning gains (section 4.4.1.). However, it is as yet unclear whether authentic blog posts and vlogs also frequently repeat low-frequency words, which lend themselves to incidental learning. This could be investigated in future corpus studies like those that have been conducted with movies (Webb, 2010) and television shows (Rodgers & Webb, 2011; Sockett, 2011a; Webb & Rodgers, 2009). If vlogs and blogs do not repeat low-frequency words as frequently as the pseudo-words in the present study, the gains in the present study would be an overestimation of authentic learning gains (cf. Rodgers, 2013; discussed in section 2.4.3.).

Finally, the tests used to evaluate vocabulary gains in this study may limit the generalisability of the results. For example, the participants may have gained knowledge of other words that were not tested. By selecting participants on the basis of their VLT scores, it was assumed that they were able to recognise the meaning of the most frequent 5,000 words in English, or around 96% of the words in the input materials. Nevertheless, encountering vocabulary in context, such as in a text or video, provides a strong sense of word meaning and use (Huckin & Coady, 1999). The treatment may therefore have led participants to increase the depth of their knowledge of words whose meanings were already known to them, gaining, for example, knowledge of the phonological forms, alternate meanings, or collocations.

This section discussed a number of factors, which may have led to under- or overestimation of incidental learning gains in this study. Additionally, previous research (section 2.5.) has identified a host of other variables that affect incidental vocabulary acquisition, which were not controlled in this study. Because of a lack of strong evidence to suggest the contrary it could be assumed that these affected both the video and blog condition equally (however, see Vidal, 2011; described in section 2.4.2.), so that the differences found between them could be generalisable to other blogs and vlogs. It is less clear to what extent the findings are generalisable to other text and video materials, since each genre of language input brings with it its own characteristics.

6.4. Research Question 3: Interactions with the Input Media

The third research question asked how learners made use of the unique features available to them on the social media platforms.

6.4.1. User-user interaction

According to Toffoli and Sockett (2010) and Woo and Reeves (2007), one of the advantages of online media is the possibility for meaningful interaction with other users, which according to the Interaction Hypothesis (Long, 1981) facilitates language learning. The screen-captures indicated that participants in both groups were likely to skim other users' comments on the blogs and videos but only nine (five in the blog, four in the video group) decided to leave comments themselves.

While this may seem like a low rate of participation, it must be noted that indeed 11% of the study's participants decided to leave a comment. From the limited data available, this seems to be representative, or even an overestimate, of how often users decide to engage with blog posts and video blogs at large: in a survey of more than 100,000 members of the *Nerdfighter* YouTube community, the majority of respondents (60%) reported that they 'never' or 'rarely' comment on YouTube videos while less than 4% reported commenting on most or every video (Green, 2014). Furthermore, 29% of respondents reported that they commented only if they felt that they had something very important to say (ibid.). The fact that participants expressed more interest in the blog posts than the videos (section 5.5) may explain the imbalance in the number of comments left on both platforms.

Finally, it is worth noting that the comments on the blog posts included eight instances of participants using the pseudo-words while they were not used in any comments on the video blogs. This might be because the participants in the video group may have been uncertain about the pseudo-words' orthographic forms, seeing as they also showed fewer gains in this aspect of knowledge (section 6.2.).

6.4.2. User-content interaction

The screen-captures were collected mainly to investigate the participants' interactions with the input materials themselves. Possible interactions with the blog posts were much like those with any other text: learners could skim the text or skip back and forth (pre- or re-read certain words or sections) to help them comprehend the overall message and infer the meaning of unknown words. As mentioned before (section 6.1.), the fact that participants scrolled through the blogs at widely varying paces suggests that some may have processed them more deeply than others. Otherwise, the screen-captures revealed only few instances of participants skipping back and forth in the blogs and when this did happen, it was clear that they were scrolling to look at the photographs referred to in the text. This could itself be seen as a comprehension strategy, trying to use the available visual material to comprehend the written meaning, and lends support to the Multiple-channel Theory (Moore et al., 2004), which states that processing information in more than one form (e.g. image and text) is helpful to the learner.

In the video group, participants would have been able to pause the video, skip ahead, or re-wind. Additionally, they had the option of turning on English captions provided with each video. The two participants in the video group who submitted screen-captures, however, showed no instances of using these strategies. It must be taken into consideration, however, that participants may not have noticed that captions were available since most vlogs provide only less-than-accurate *auto-caption*s (automatically generated by YouTube's voice recognition software). Non-automatic captions are only available where the creator manually adds them.

The fact that the participants in both groups did not make full use of the options for engagement available to them could indicate that they felt they had already sufficiently comprehended the input. It must be remembered that the participants' general vocabulary knowledge was pre-tested and the videos' lexical content analysed to ensure that only a small number of words (about 4%) would be unknown to them. Perhaps the learners would have used these strategies if they had had difficulty inferring the meaning of more words or understanding the overall message.

Nevertheless, it could be hypothesised that, if the participants had taken full advantage of the features available to them, they may have shown even greater learning gains. For example, if participants in the video group had chosen to use the subtitles, they may have gained more orthographic knowledge.

At this point, it is worth discussing the usefulness of the screen-captures on which part of the analysis in this section rests. Although some inferences could be made from them, they did not prove as useful as it had previously been thought. They were an attempt at creating a more objective way of collecting usage data than through self-reporting (cf. Sockett & Toffoli, 2012; Toffoli & Sockett, 2010). However, only few screen-captures were collected and they included less information than had been hoped.

Because recording and uploading the screen-captures was time-intensive, it was foreseen that only few participants would submit them and the option to partake in the study without recording a screen-capture was added. However, the large imbalance between the number of participants who submitted screen-captures in the blog and video groups (ten and two respectively) was surprising.

Since the web survey's randomisation algorithm was instructed to assign equal numbers of participants to the video and blog groups, the lower number of responses in the video screen-capture group must be due to a higher dropout rate. In fact, eleven participants completed this condition, but only two included links to the screen-captures. Some of the others commented that they were unable to share the videos, as their uploads were too slow or failed completely. That this was a problem in the video but not the blog group can likely be attributed to the fact that a screen-capture of watching a video produces a much larger file than a screen-capture of reading a blog post, therefore requiring more time and bandwidth to upload and being more prone to failure. In this regard, this part of the data collection failed due to technical issues, which had not come to light during the pilot testing. If these could be solved, a follow-up study with a bigger sample may provide valuable insights into the way in which users interact with, and can learn from, social media platforms.

Another issue with the screen-captures was that some behaviours could not be clearly discerned, for example the speed at which the learners read the

blog posts, or which user comments they attended to. Participants who volunteered to record a screen-capture were given instructional videos modelling how participants could clearly show how they were interacting with the websites, e.g. using the mouse cursor to indicate what they were reading. However, only two out of the twelve participants who submitted screen-captures actually used these strategies. The others could have felt that this behaviour was too unnatural or they might not have watched the instructional video. Future studies may find a way to ensure that all participants watch the tutorial.

All in all, while some information could be discerned from the screen-captures, a questionnaire about the participants' interactions may have been more suited researching the way in which the participants engaged with the blogs and video blogs. Alternatively, this research question could be addressed in more detail in future self-report studies, like those conducted by Sockett and Toffoli (2012; Toffoli & Sockett, 2010).

6.5. Research Question 4: Appeal and Familiarity of Input Materials

The last research question asked to what extent the social media used in this study appealed to the participants and to what extent they were they similar to those that users would choose to engage with independently. As might be expected, the participants reported different levels of liking the input; in fact, this variable was roughly normally distributed in both groups (Figure 4, section 5.5.). Interestingly, however, more participants seemed to be interested in the topic of online friendships when it was presented to them in the form of a blog (Figure 5, section 5.5.), implying that there must be some conflation of the topic and modality in these responses. As the format of the blog posts was also less familiar to the participants than that of the vlogs (Figure 6, section 5.5.), it is conceivable that this feeling of novelty is reflected in the blog group's heightened levels of interest.

The different levels of familiarity could suggest better ecological validity of the findings in the video group. It may have arisen because the blogs were created from the videos rather than vice versa. An alternative explanation could be that the participants had different conceptions of a *blog* than the researcher: The blog posts in this study were relatively long (between 450 and 800 words) and text-heavy, similar to posts that may be found on popular sites like Blogspot and

WordPress. However, other *blogging* platforms, such as Tumblr and Pinterest mostly contain short text posts (under 200 words), pictures, and even videos. If the participants' conception of a blog was based on websites like these, it would be unsurprising that the format of the posts in this study were less familiar to them.

Finally, it was asked to that extent learners habitually engage with the types of media used in this study. Most participants reported engaging with blogs or video blogs either daily (60% in the video and 81% in the blog group) or two to three times per week (30% and 12% respectively). All participants also reported subscribing to at least one blog or YouTube channel; most indicated subscriptions to between one and ten blogs or channels (61% and 71% of participants respectively). Only few stated that they subscribed to more than 50, although this was more common in the blog than the video group (20% versus 5% of participants). This discrepancy could, once again, arise from a different conception of what a blog is, since on platforms like Tumblr and Pinterest, where posts often consist of very short texts or images, it is not uncommon for users to subscribe to large numbers of blogs. It must be noted that participants in each group were only asked about their engagement with the input medium they had been exposed to in the study. Upon reflection, it would have been interesting to ask all participants about their engagement with both media. This would have been another way of testing the underlying similarity or differences between the two groups.

Altogether, though, these numbers speak to the large potential of these media to contribute to young peoples' Online Informal Learning of English. However, it must be kept in mind that the participants in this study were sampled from a virtual community, which exists primarily on social networking sites, blogs, and vlogs. It is unclear to what extent the results of this study are then generalisable to users who are not members of virtual communities. Previous research with less specialised populations of ESL learners (Sockett & Toffoli, 2012; Toffoli & Sockett, 2010) does not mention interactions with blogs or vlogs, so the frequency of engagement cannot be compared.

6.6. Limitations and Suggestions for Future Research

Some limitations of this study have been discussed throughout this chapter, such as the generalisability, sample size, the use of pseudo-words, and the fact that the in-text frequency of the target words was controlled.

Furthermore, it must be noted that vocabulary gains were measured based on only six target words (three nouns, two verbs, one adjective). Future research should compare incidental learning video of more target words through reading and watching videos and could also include other parts of speech. Furthermore, only three aspects of vocabulary knowledge (grammatical function, spelling, meaning) were assessed. However, using a more extensive battery of tests in future studies would enable researchers to compare in more detail which aspects of vocabulary are best learnt from which type of input. Due to the study's limited time frame and method of data collection, no delayed post-tests were conducted. However, future studies should compare the durability of learning gains from the two input modalities. Finally, although the present study controlled for some factors that have previously been found to influence incidental vocabulary learning (vocabulary size, in-text frequency of target words), it did not account for a number of other variables, such as the contextual informativeness or the learners' L2 proficiency. It would be useful to explore in future research whether there are differences in the extent to which these factors affect vocabulary learning through different social media.

Chapter 7: Conclusion

This study adds to the existing body of research on incidental vocabulary learning in several ways. Inspired by research on Online Informal Learning of English (Sockett & Toffoli, 2012; Toffoli & Sockett, 2010), it was (to the best of the author's knowledge) the first to compare incidental learning gains from two popular types of online social media: blog posts and video blogs. Unlike the majority of the texts and videos studied in previous research (graded readers, academic texts and lectures, documentaries, educational videos), the media investigated here are not usually intended as learning resources. Therein, they may be more representative of content which L2 users would likely choose to engage with independently. By directly comparing the learning gains from both media, the present study furthermore contributes to the still very small body of research (Neuman & Koskinen, 1992; Vidal, 2011) on the differential effects of reading versus watching video on incidental vocabulary learning.

The results showed that incidental vocabulary learning occurred, in approximately equal amounts, from reading blog posts and watching vlogs. However, different types of vocabulary knowledge were gained from the two modalities. Unsurprisingly, written input promoted greater gains in orthographic knowledge than videos, although previous research findings (e.g. Hui, 2007; Markham, 1999; Winke et al., 2010) indicate that this effect would likely shrink or disappear if the participants had watched the vlogs with L2 captions. In turn, the videos were found to promote greater recognition and recall of the target words' meaning and recall of their grammatical function. This may be because vlogs are highly engaging and the visual information may help learners decode the meaning of the aural language input (Moore et al., 2004).

There seems to be a commonly held belief among language professionals that extensive reading is the most important way of learning vocabulary (e.g. Waring, 2009). In light of the present findings, however, it seems a serious oversight to discount the effectiveness of social media for word learning. In OILE, the users themselves choose which materials to engage with, which presumably leads to higher levels of motivation and allows for individualisation of the vocabulary learnt. In a language class that includes students with widely varying interests, it is often impossible to teach every individual the vocabulary he or she needs. However, a student interested in mathematics might, for example, choose

to follow YouTube video blogger Vi Hart (2014), who produces highly entertaining educational videos explaining mathematical theories and principles. Based on the findings of this study, her videos seem to have the potential to teach our student the technical vocabulary she needs to pursue her interest in mathematics further.

Even though students engage in OILE independently, knowing about the different types of lexical knowledge that are promoted by various input materials can furthermore help teachers and learners to find the most effective methods of teaching and learning to reach a given goal. Since the analysis of participants' interactions with the blogs and vlogs showed that they did not take full advantage of the opportunities for engagement offered by these media, teachers might also consider addressing the potential of OILE in their language classes and teaching strategies to maximise learning from social media platforms. It should be taken into consideration, however, that discussing online activities as learning tasks could negatively impact students' motivation to partake in these activities in the future.

References

Adolphs, S. (2003). Lexical Coverage of Spoken Discourse. *Applied Linguistics*, *24*(4), 425–438. doi:10.1093/applin/24.4.425

Adolphs, S., & Schmitt, N. (2004). Vocabulary coverage according to spoken discourse context. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language: Selection, acquisition, and testing* (pp. 39–49). Amsterdam: John Benjamins.

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in Reading and How Children Spend Their Time Outside of School. *Reading Research Quarterly*, *23*(3), 285–303. doi:10.1598/RRQ.23.3.2

Ault, S. (2014). Survey: YouTube Stars More Popular Than Mainstream Celebs Among U.S. Teens. *Variety*. Retrieved August 05, 2014, from http://variety.com/2014/digital/news/survey-youtube-stars-more-popular-than-mainstream-celebs-among-u-s-teens-1201275245/

Blattner, G., & Fiori, M. (2009). Facebook in the Language Classroom: Promises and Possibilities. *International Journal of Instructional Technology & Distance Learning*, *6*(1), 17–28. Retrieved from http://itdl.org/journal/jan_09/article02.htm

Blattner, G., & Lomicka, L. (2012). Facebook-ing and the Social Generation: A New Era of Language Learning. *Alsic*, *15*(1), 1–32. doi:10.4000/alsic.2413

Bonk, W. J. (2000). Second Language Lexical Knowledge and Listening Comprehension. *International Journal of Listening*, *14*(1), 14–31. doi:10.1080/10904018.2000.10499033

Bruton, A., López, M. G., & Mesa, R. E. (2011). Incidental L2 Vocabulary Learning: An Impracticable Term? *TESOL Quarterly*, *45*(4), 759–768. doi:10.5054/tq.2011.268061

Brysbaert, M., & New, B. (2009). SUBTLEX-US Word Frequency American English. *Department of Experimental Psychology, University of Gent*. Retrieved May 07, 2014, from http://expsy.ugent.be/subtlexus/

Cameron, A. C. (2004). Kurtosis. In M. S. Lewis-Beck, A. Bryman, & T. F. Liao (Eds.), *The SAGE Encyclopedia of Social Science Research Methods* (pp. 544–546). Thousand Oaks, CA: Sage. Retrieved from http://srmo.sagepub.com/view/the-sage-encyclopedia-of-social-science-research-methods/n467.xml

Çetin, Y., Sözcü, Ö. F., & Kinay, H. (2012). Incidental foreign language vocabulary acquisition from social network games. *Incidental Foreign Language Vocabulary Acquisition from Social Network Games*, *9*(2), 535–552. Retrieved from http://www.j-humansciences.com/ojs/index.php/IJHS/article/view/2223

Cobb, T. (2013). Lextutor Vocab Profiler (BNC-COCA). *Compleat Lexical Tutor v.6.2*. Retrieved July 17, 2014, from http://www.lextutor.ca/vp/bnc/

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. doi:10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of Processing and the Retention of Words in Episodic Memory. *Journal of Experimental Psychology: General*, *4*(3), 268–294. doi:10.1037.0096-3445.104.3.268

Davies, M. (2014). Copus of Global Web-Based English. *Brigham Young UNiversity*. Retrieved May 09, 2014, from http://corpus2.byu.edu/glowbe/

DeKeyser, R. M., & Sokalski, K. J. (1996). The Differential Role of Comprehension and Production Practice. *Language Learning*, *46*(4), 613–642. doi:10.1111/j.1467-1770.1996.tb01354.x

Elley, W. B. (1991). Acquiring Literacy in a Second Language: The Effect of Book-Based Programs *. *Language Learning*, *41*(3), 375–411. doi:10.1111/j.1467-1770.1991.tb00611.x

Ellis, N., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology Section A*, *46*(3), 533–558. doi:10.1080/14640749308401062

Ellis, R. (1999). Factors in the Incidental Acquisition of Second Language Vocabulary from Oral Input. In R. Ellis (Ed.), *Learning a Second Language through Interaction* (pp. 35–62). Amsterdam: John Benjamins.

Ethical Guidelines for Educational Research. (2011). *Association, British Educational Research*. Retrieved August 16, 2014, from http://www.bera.ac.uk/wp-content/uploads/2013/12/BERA-Ethical-Guidelines-2011.pdf

Gass, S. M. (1999a). Discussion: Incidental Vocabulary Learning. *Studies in Second Language Acquisition*, *21*(2), 319–333. doi:10.1017/S0272263199002090

Gass, S. M. (1999b). *Input, Interaction, and the Second Language Learner*. Oxford: Routledge.

Green, H. (2014). 2014 Nerdfighter Census Analysis. Retrieved June 10, 2014, from https://www.youtube.com/watch?v=Sh5RemPWcds

Harley, B. (1995). The lexicon in language research. In *Lexical issues in language learning* (pp. 1–28). Amsterdam: John Benjamins.

Henriksen, B. (1999). Three Dimensions of Vocabulary Development. *Studies in Second Language Acquisition*, *21*(02), 303–317. doi:10.1017/S0272263199002089

Henriksen, B., Albrechtsen, D., & Haastrup, K. (2004). The Relationship between Vocabulary Size and Reading Comprehension in the L2. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Angles on the English-Speaking World: Writing and Vocabulary in Foreign Language Acquisition* (pp. 129–140). Copenhagen: Museum Tusculanum Press.

Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, *8*(2), 689–696. Retrieved from http://nflrc.hawaii.edu/rfl/PastIssues/rfl82hirsh.pdf

Horst, M. (2005). Learning L2 Vocabulary through Extensive Reading: A Measurement Study. *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes*, *61*(3), 355–382. doi:10.3138/cmlr.61.3.355

Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring Second Language Vocabulary through Reading. *Reading in a Foreign Language*, *11*(2), 207–223. Retrieved from http://nflrc.hawaii.edu/rfl/PastIssues/rfl112horst.pdf

Hu, M. H.-C., & Nation, I. S. P. (2000). Unknown Vocabulary Density and Reading Comprehension. *Reading in a Foreign Language*, *13*(1), 403–430. Retrieved from http://nflrc.hawaii.edu/rfl/PastIssues/rfl131hsuehchao.pdf

Huang, B.-G., & Yang, J. C. (2012). A Multiplayer Online Role-Playing Game for Incidental Vocabulary Learning. In *Proceedings of the 20th International Conference on Computers in Education*. Singapore: National Institute of Education, Nanyang Technological University. Retrieved from http://www.lsl.nie.edu.sg/icce2012/wp-content/uploads/2012/12/C6-f-194.pdf

Huckin, T., & Coady, J. (1999). Incidental Vocabulary Acquisition in a Second Language. *Studies in Second Language Acquisition*, *21*(2), 181–193. doi:10.1017/S0272263199002028

Hui, W. (2007). The effects of captions on Chinese EFL students' incidental vocabulary acquisition. *Chinese English Language Education Association Journal*, *30*(4), 9–16. Retrieved from http://www.celea.org.cn/teic/74/74-09.pdf

Hulstijn, J. H. (1990). A comparison between the information-processing and the analysis/control approaches to language learning. *Applied Linguistics*, *11*(1), 30–45. doi:10.1093/applin/11.1.30

Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: a reappraisal of elaboration, rehearsal and automaticity. In P.

Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 258–286). Cambridge: Cambridge University Press.

International Friendships. (2014). *blogspot.com*. Retrieved May 15, 2014, from http://oxeddiss.blogspot.co.uk

Jing. (2014). *TechSmith Corporation*. Retrieved July 04, 2014, from http://www.techsmith.com/download/jing/

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford; New York: Pergamon Press.

Kucera, H., & Francis, W. N. (1967). *A computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kweon, S.-O., & Kim, H.-R. (2008). Beyond raw frequency: Incidental vocabulary acquisition in extensive reading. *Reading in a Foreign Language*, *20*(2), 191–215. doi:1539-0578

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special Language: From Humans Thinking to Thinking Machines* (pp. 316–323). Clevedon: Multilingual Matters.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 20–34). Cambridge: Cambridge University Press.

Laufer, B. (1998). The Development of Passive and Active Vocabulary in a Second Language: Same or Different? *Applied Linguistics*, *19*(2), 255–271. doi:10.1093/applin/19.2.255

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1–26. doi:10.1093/applin/22.1.1

Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*, 33–51. doi:10.1177/026553229901600103

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: effects of language learning context. *Language Learning*, *48*(3), 365–391. doi:10.1111/0023-8333.00046

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited : Lexical text coverage , learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, *22*(1), 15–30. Retrieved from http://nflrc.hawaii.edu/rfl/April2010/articles/laufer.pdf

Liu, N., & Nation, I. S. P. (1985). Factors Affecting Guessing Vocabulary in Context. *RELC Journal*, *16*(1), 33–42. doi:10.1177/003368828501600103

Lomicka, L., & Lord, G. (2009). *The Next Generation: Social Networking and Online Collaboration in Foreign Language Learning*. San Marcos, TX: The Computer Assisted Language Instruction Consortium. Retrieved from https://calico.org/page.php?id=430

Long, M. H. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Science*, *379*(1), 259–278. doi:10.1111/j.1749-6632.1981.tb42014.x

Markham, P. (1999). Captioned Videotapes and Second-Language Listening Word Recognition. *Foreign Language Annals*, *32*(3), 321–328. doi:10.1111/j.1944-9720.1999.tb01344.x

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary - Description, Acquisition and Pedagogy* (pp. 84–102). Cambridge: Cambridge University Press.

Milton, J. (2008). Vocabulary uptake from informal learning tasks. *Language Learning Journal*, *36*(2), 227–237. doi:10.1080/09571730802390742

Milton, J., Wade, J., & Hopkins, N. (2010). Aural Word Recognition and Oral Competence in English as a Foreign Langguage. In R. Chacón-Beltrán, C. Albello-Contesse, & M. Torreblanca-López (Eds.), *Insights into Non-native Vocabulary Teaching and Learning* (pp. 83–98). Bristol: Multilingual Matters.

Moore, M., Burton, J. K., & Myers, R. J. (2004). Multiple-channel communication: The theoretical and research foundations of multimedia. In D. H. Jonasssen (Ed.), *Handbook of research on educational communications and technology* (2nd ed., pp. 979–1005). Mahwah, NJ: Lawrence Erlbaum.

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19–35). Hillsdale, NK: Lawrence Erlbaum.

Nation, I. S. P. (1983). Testing and Teaching Vocabulary. *Guidelines*, *5*, 12–25.

Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Boston, MA: Heinle.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* (C. A. Chapelle & S. Hunston, Eds.). Cambridge: Cambridge University Press.

Nation, I. S. P. (2006). How Large a Vocabulary is Needed For Reading and Listening ? *The Canadian Modern Language Review*, *63*(1), 59–82. doi:10.1353/cml.2006.0049

Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary - Description, Acquisition and Pedagogy* (pp. 6–19). Cambridge: Cambridge University Press.

Neuman, S. B., & Koskinen, P. (1992). Captioned Television as Comprehensible Input: Effects of Incidental Word Learning from Context for Language Minority Students. *Reading Research Quarterly*, *27*(1), 94–106. doi:10.2307/747835

Norman, D. A. (1976). *Memory and attention: an introduction to human information processing*. London: John Wiley.

OxEd. (2014). *youtube.com*. Retrieved May 15, 2014, from https://www.youtube.com/channel/UCoFpDjwk5c-jBbbYVEJMgmQ

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*(3), 331–347. doi:10.1016/0749-59X(91)90040-Q

Paribakht, T. S., & Bingham Wesche, M. (1993). Reading Comprehension and Second Language Development in a Comprehension-Based ESL Program. *TESL Canada Journal*, *11*(1), 9–29. Retrieved from http://www.teslcanadajournal.ca/index.php/tesl/article/view/623

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do "Things Fall Apart"? *Reading in a Foreign Language*, *22*(1), 31–55. Retrieved from http://nflrc.hawaii.edu/rfl/April2010/articles/pellicersanchez.pdf

Peters, E. (2006). L2 vocabulary acquisition and reading comprehension: The influence of task complexity. In M. P. García Mayo (Ed.), *Task in Formal Language Learning* (pp. 178–198). Clevedon: Multilingual Matters.

Peters, E., Hulstijn, J. H., Sercu, L., & Lutjeharms, M. (2009). Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared. *Language Learning*, *59*(1), 113–151. doi:10.1111/j.1467-9922.2009.00502.x

Pigada, M., & Schmitt, N. (2006). Vocabulary Acquisition from Extensive Reading: A Case Study. *Reading in a Foreign Language*, *18*(1), 1–28. Retrieved from http://nflrc.hawaii.edu/rfl/April2006/pigada/pigada.pdf

Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics*, *26*, 66–86. doi:10.1093/applin/aml049

Qualtrics Research Suite. (2014). *Qualtrics, LLC*. Retrieved May 08, 2014, from http://www.qualtrics.com

Rodgers, M. P. H. (2013). *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions*. Unpublished PHD Dissertation. Victoria University of Wellington.

Rodgers, M. P. H., & Webb, S. (2011). Narrow Viewing: The Vocabulary in Related Television Programs. *TESOL Quarterly*, *45*(4), 689–717. doi:10.5054/tq.2011.268062

Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, *6*(2), 72–78. doi:10.1016/0346251X(78)90027-1

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158. doi:10.1093/applin/11.2.129

Schmitt, N. (2000). *Vocabulary in Language Teaching*. (J. C. Richards, Ed.). Cambridge: Cambridge University Press.

Schmitt, N. (2010). *Researching vocabulary: a vocabulary research manual*. Basingstoke: Palgrave MacMillan.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*(1), 55–88. doi:10.1177/026553220101800103

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *45*(1), 21–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1636010

Social Networking Sites. (2012). *Get Safe Online*. Retrieved April 17, 2014, from https://www.getsafeonline.org/social-networking/social-networking-sites/

Sockett, G. (2011a). From the cultural hegemony of English to online informal learning: Cluster frequency as an indicator of relevance in authentic documents. *ASp*, *59*, 5–20. doi:10.4000/asp.2210

Sockett, G. (2011b). Les processus cognitifs de résolution de problèmes pour l'apprentissage des langues dans des encironnements multimédias: Apprentissage informel et résaux sociaux. *Les Cahiers de l'Acedle*, *8*(1), 29–46. Retrieved from http://acedle.org/spip.php?article3184

Sockett, G. (2012a). L'impact des activités informelles d'écoute en anglais sur les travaux écrits d'étudiants LanSAD, une analyse quantitative. In *Colloque 2012 de l'Association des Chercheurs et Enseignants en didactique des langues (acedle) - Apprendre les languages autrement*. Nantes, France.

Sockett, G. (2012b). Le web social – La complexité au service de l' apprentissage informel de l' anglais. *Alsic*, *15*(2). doi:10.4000/alsic.2505

Sockett, G., & Toffoli, D. (2012). Beyond learner autonomy: a dynamic systems view of the informal learning of English in virtual online communities. *ReCALL*, *24*(02), 138–151. doi:10.1017/S0958344012000031

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139–152. doi:10.1080/09571730802389975

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. 235–252). Rowley, MA: Newbury House.

Thorne, S., & Reinhardt, J. (2008). "Bridging activities", New Media Literacies and Advanced Foreign Language Proficiency. *CALICO Journal*, *25*(3), 558–572. Retrieved from https://calico.org/html/article_717.pdf

Toffoli, D., & Sockett, G. (2010). How non-specialist students of English practice informal learning using web 2.0 tools. *ASp*. doi:10.4000/asp.1851

Toffoli, D., & Sockett, G. (2013). University teachers' perceptions of Online Informal Learning of English (OILE). *Computer Assisted Language Learning*, (October), 1–15. doi:10.1080/09588221.2013.776970

Vi Hart. (2014). *youtube.com*. Retrieved August 22, 2014, from https://www.youtube.com/user/Vihart

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, *24*(1), 56–89. doi:10.1093/applin/24.1.56

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219–258. doi:10.1111/j.1467-9922.2010.00593.x

vlogbrothers. (2014). *youtube.com*. Retrieved May 09, 2014, from https://www.youtube.com/user/vlogbrothers?gl=GB&hl=en-GB

Waring, R. (2009). The inescapable case for extensive reading. *A. Cirocki (Ed.) Extensive Reading in English Language Teaching*. Retrieved August 18, 2014, from http://www.robwaring.org/er/what_and_why/er_is_vital.htm

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*(2), 130–163. Retrieved from http://nflrc.hawaii.edu/rfl/October2003/waring/waring.pdf

Webb, S. (2007). The Effects of Repetition on Vocabulary Knowledge. *Applied Linguistics*, *28*(1), 46–65. doi:10.1093/applin/aml048

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, *20*(2), 232–245. Retrieved from http://nflrc.hawaii.edu/rfl/October2008/webb/webb.pdf

Webb, S. (2010). A corpus driven study of the potential for vocabulary learning through watching movies. *International Journal of Corpus Linguistics*, *15*(4), 497–519. doi:10.1075/ijcl.15.4.03web

Webb, S., & Rodgers, M. P. H. (2009). Vocabulary Demands of Television Programs. *Language Learning*, *59*(2), 335–366. doi:0023-8333

Wesche, M., & Paribakht, T. S. (1996). Assessing vocabulary knowledge: depth vs. breadth. *Canadian Modern Language Review*, *53*(1), 13–49.

Winke, P., Gass, S. M., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, *14*(1), 65–86. Retrieved from http://llt.msu.edu/vol14num1/winkegasssydorenko.pdf

Woo, Y., & Reeves, T. C. (2007). Meaningful interaction in web-based learning: A social constructivist interpretation. *The Internet and Higher Education*, *10*(1), 15–25. doi:10.1016/j.iheduc.2006.10.005

Yuksel, D., & Tanriverdi, B. (2009). Effects of watching captioned movie clip on vocabulary development of EFL learners. *The Turkish Online Journal of Educational Technology*, *8*(2), 48–54. doi:1303-6521

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring Vocabulary through Reading: Effects of Frequency and Contextual Richness. *The Canadian Modern Language Review*, *57*(4), 541–572. Retrieved from http://www.lextutor.ca/cv/v_conditions.htm

Appendices

Appendix 1: Input Materials: The three blog posts

1.1. Chris on Making YouTube friends

Hey guys!
A lot of you have been asking me about the sagods I have on my wall so I thought I'd tell you the story behind them. I'll start with one sagod in particular and it's this one of me and JP: [photograph removed].

So I think it was around April 2012 and I'd just uploaded a big video about Doctor Who on my channel. And I get this masco from a YouTuber in Australia saying we should ancon a video together and collaborate. So I ancon a reply saying it's a really cool idea, it's just a shame that he's on the other side of the world! But he ancons back – surprisingly quickly actually considering the different pacons – and he says he's planning to tasper around Europe and we should meet up. So I just thought 'Why not?' and he taspers to the UK I think in June the same year, and he sends me a masco saying 'Let's go for sushi!'.

So we met up in London and I was really nervous actually because it was the first time I'd just ever agreed to just go and meet a YouTube who I've never met before. But it was really dangy actually and we hit it off straight away. So we shot a little vid together about the differences between the UK and Australia and there's quite a lot of differences actually. There's more than just the accents and the different pacons. And I find it really dangy actually that Australians love coming to the UK so much. I mean, why would you want to tasper over here when you've got the amazing beaches over there?

But yeah, we anconed this little script together, made our little video, and after that he carried on taspering. I think he went off to Germany and I thought that was the last we'd see of each other. But then I was at Summer in the City a few weeks later and I got a masco saying "Hey! I'm coming to Summer in the City, too!".

So we hung out that weekend and I introduced him to all my friends and that's actually when we took that sagod. We also took another one with everyone: [photograph removed].

And I think it's just really cool that you can meet people from the other side of the world through YouTube. And it's just really dangy that it didn't exist a few years ago and then here it is, creating friendships!

But yeah, even though he's back in Australia we still send each other mascos and we always try and Skype whenever we can. Obviously there's a bit of a difference in the pacons so we have to see when we're both online. But whenever we can, we do. I think that's really cool and that's why I keep all of the sagods of me and my friends on my wall.

1.2. Lidewij on Online Friendships

Hi guys! In this post I want to talk about one of my favorite benefits of being involved in the YouTube community: Online friendships.

I think online friendships are great and can be pretty dangy as well. I highly recommend acquiring some! I'll try to add sagods whenever I have any, but no promises!

My first real live contact with people I knew through the internet was in Berlin, in 2008. John Green, who is a YouTuber and writer, had just anconed Paper Towns and was visiting the city for some sort of book event. Approximately six people who watched his videos taspered to Berlin to see him. Here, a sagod: [photograph removed].

It was very difficult to get time off from school to tasper to Berlin, but it worked out in the end. The most dangy take-away was, in my opinion, how these people I met, despite growing up in a different country, had similar inside jokes and stuff. It was the first time I realized that being part of an internet community can have such a profound impact on your opinions, humor and culture. I think that's really dangy!

After I got back, I anconed a masco in response to a thread in which people were trying to organize a gathering in Amsterdam. We met up pretty soon after. Six of the people I met at that gathering are still some of my best friends. If I manage to find a sagod, I'll put it right here: [photograph removed].

In 2010 I went to VidCon, which is a large YouTube event in LA. This was the first gathering I attended outside of my own pacon. I finally got to meet some people I'd only known online before, specifically my Canadian friend Jess. This is a sagod of us in LA: [photograph removed].

After VidCon I sent Jess a masco to ask if I could visit her in Canada and she sent a masco back saying it was OK. I had a great time staying with her and learned some dangy stuff about Canada, like that you can't boo people at sporting events and that home-made poutine is *not* a thing.

I guess there are some downsides to having internet friends as well. Taspering can be fun but it can also suck. My trip back to Europe from Canada was the *worst*! I very, very nearly missed my flight and then I didn't feel well during the flight at all. And finally the transition from an American pacon to a European pacon is always particularly difficult.

Another obvious downside of having friends on the other side of the world is that you can only ancon to them but don't get to see them often. Which is not cool. If someone is eight or nine pacons away from you it's sometimes difficult to even find a time to talk. On a happier note though: Jess just sent me a masco to let me know that she is taspering to Europe soon! Wohooo! Very exciting!

I'm interested to hear what you think about this topic. Leave me a comment below, or send me a masco! Ciao!

1.3. JP on International Friendships

Hello! It's been a little while. We don't speak often enough. Which is dangy because that's what I was going to talk about today. Awesome segway...

I find myself not anconing friends enough, especially to my international friends. I've done a lot of taspering and thus have made a lot of friends in foreign countries. Some of these people I have met at gatherings like Summer in the City. Cue sagod: [photograph removed].

A great place to make friends. And lose them. In the crowd!

Being part of the YouTube community is so great! It's just somewhere you meet like-minded people and gathering**s** are a great way of bringing us all together. It has gained me many close friends but, unfortunately, a lot of these people live in countries other than Australia.

Some of the people I've met at international events are some of my closest friends. And yet still, somehow, I find remembering to message them challenging. I think this problem has a lot to do with convenience. Like, it's easy just to keep in contact with those around you but remembering to keep in contact with people in other countries can be rarer than you'd expect. It's funny and also somewhat dangy that this would still happen these days considering we have so many convenient ways of sending mascos**.** Especially to our international friends.

But there are other factors you see? You don't see? Oh, well I'll show you!

You see, being in different pacons really slows down the conversation when you're anconing and it's hard to organize Skype calls when you're constantly sleeping and working at different times. I suppose we end up connecting to these friends through other mediums like watching their videos or seeing their sagods on instagram and Facebook. Sometimes it feels like like we're connecting to them by liking their stuff and all when really we're forgetting to actually send mascos to them. A lot of the time I don't masco these people because I feel I have nothing dangy to tell them. So that's also a major problem in my opinion.

I wish I could tasper all the time because I miss these friends a lot. And sometimes videos and sagods just don't cut it. Well... Unless it's a paper cut. Wait. Do people even print sagods any more?

Anyway, despite being in different pacons and different countries I think it's important that we make sure to remember to keep anconing these friends. Because, essentially, communication is the most important part of an international relationship. Even short mascos can do a lot for a friendship. Yep. Bit of a wise masco there from my ancient brain. Very dangy... I think YouTube gatherings are a great way of getting to see these people again though. There's usually a few per year and it can bring you together with many friends at once.

Although there's problems with that. You see? You don't see? Okay.

Firstly, I'm not in the right country. Most of the gatherings are in America and

England and I'm nowhere near either of those countries. Even though these gatherings are very international events, it's usually only the popular YouTubers who get to tasper there**.** I was planning on going to Summer in the City again this year but… I missed out on tickets. I have pacons to blame because the tickets went on sale when I was sleeping and I didn't remember to stay up to buy them. Boo hoo. Screw pacons! To be honest I don't really have the money to tasper right now anyway. Haha, broke life! I'll just have to enjoy the sagods of it.

Even though these things can be difficult, I intend to tasper to many future gatherings**.** And to keep anconing because the people are amazing and so worth it! Just a simple masco, shared video or sagod and we can keep these international friendships going. As we should. And that's what I'm gonna do.

Thanks for reading! I hope this post was dangy for you. I'm sure many of you could relate.

After all, this is the internet! Peace and Love in extremities! Bye!

Appendix 2: Dependent Measures

2.1. Orthography

In each of the sound files in this section you will hear a word read out to you twice. Please write each word in the text field below the audio player. If you have no idea how to write a word, do not guess. In that case, you can write 'I don't know' in the text field. But, if you think you might know how to spell it, plase try to find the answer.

When you're finished, you can click below to continue to the next part of the study.

Example:

Item 1:

Item 2:

[...]

2.2. Recall of Meaning

Please use the text fields to provide a short description of the meanings of the words below. If you have no idea what a word means, do not guess. In that case, you can write 'I don't know' in the text field. But, if you think you might know what a word could mean, please try to find the answer.

Examples:

Deliquesce means: to turn into liquid or to melt

Sanatorium means: an institution providing medical treatment or care, a hospital

When you're finished, you can click below to continue to the next part of the study.

collab

tasper

sagod

unfortunate

pacon

ancon

dangy

poutine

masco

acquire

2.3. Recall of Grammatical Function

Please use each of the words below in a short sentence. If you have no idea about how to use a word, do not guess. In that case, you can write 'I don't know' in the text field. But if you think you might know how to use it, please try to find the answer.

Examples:

Deliquesce: I was afraid that it would deliquesce quickly.

Sanatorium: When I had my accident, they took me to the sanatorium.

collab

acquire

ancon

dangy

pacon

poutine

unfortunate

sagod

2.4. Recognition of Grammatical Function

For each of the words below, please identify the sentence in which it is used correctly.

unfortunate
☐ It is an unfortunate.
☐ He unfortunates.
☐ It is very unfortunate.
☐ I don't know.

dangy
☐ It is very dangy.
☐ This is a dangy.
☐ She dangies something.
☐ I don't know.

ancon
☐ It is an ancon.
☐ It is very ancon.
☐ He andonced.
☐ I don't know.

poutine
☐ She poutined.
☐ This is a poutine.
☐ It is very poutine.
☐ I don't know.

sagod
☐ It is very sagod.
☐ It is a sagod.
☐ She sagods something.
☐ I don't know.

masco
☐ We masco.
☐ It is very masco.
☐ This is a masco.
☐ I don't know.

collab
☐ It is a collab.
☐ He collabs.
☐ This is very collab.
☐ I don't know.

tasper
☐ It is very tasper.
☐ You taspered.
☐ This is a tasper.
☐ I don't know.

pacon
☐ She paconed.
☐ It is a pacon.
☐ It is very pacon.
☐ I don't know.

acquire
☐ It is very acquire.
☐ They acquired something.
☐ This is an acquire.
☐ I don't know.

2.5. Recognition of Meaning

Please choose the option which best describes the meaning of each word presented below.

unfortunate
☐ unforgiving
☐ unlucky
☐ unimportant
☐ I don't know.

pacon
☐ complication
☐ time zone
☐ adjustment
☐ I don't know.

acquire
☐ sell
☐ calculate
☐ get
☐ I don't know.

sagod
☐ photo
☐ birthday
☐ month
☐ I don't know.

tasper
☐ to travel
☐ to change
☐ to finish
☐ I don't know.

poutine
☐ camper
☐ fast food
☐ vendor
☐ I don't know.

dangy
☐ small
☐ clean
☐ interesting
☐ I don't know.

masco
☐ meeting
☐ message
☐ week
☐ I don't know.

ancon
☐ to move
☐ to find
☐ to write
☐ I don't know.

collab
☐ teamwork
☐ chat
☐ production
☐ I don't know.

Appendix 3: Example of Participant Questionnaire (Video Group)

Note: Once more, there were versions of this questionnaire for both experimental groups. The alternations only pertained to the mentions of the input medium, where 'watching YouTube videos', for example, was replaced with 'reading blog posts'.

You're almost done! So that I can better analyse my results, please answer the questions below.

What gender do you identify as?

What is/are your native language(s)? A native language is the language you started to learn from birth.

What is the highest level of education you have completed?
☐ Less than High School.
☐ High School
☐ Some University,
☐ Bachelor's Degree.
☐ Master's Degree
☐ Doctoral

For how many years have you been learning English?

How often do you watch YouTube videos?

How many channels are you subscribed to?

	Very				Not at all
Did you like the vlogs you just watched?	☐	☐	☐	☐	☐
How interesting did you find the topic of the vlogs?	☐	☐	☐	☐	☐
To what extent were these vlogs like those that you usually watch?	☐	☐	☐	☐	☐

Appendix 4: Histograms Showing the Distribution of Gain Scores on Five Measures of Vocabulary Knowledge

4.1. Orthography

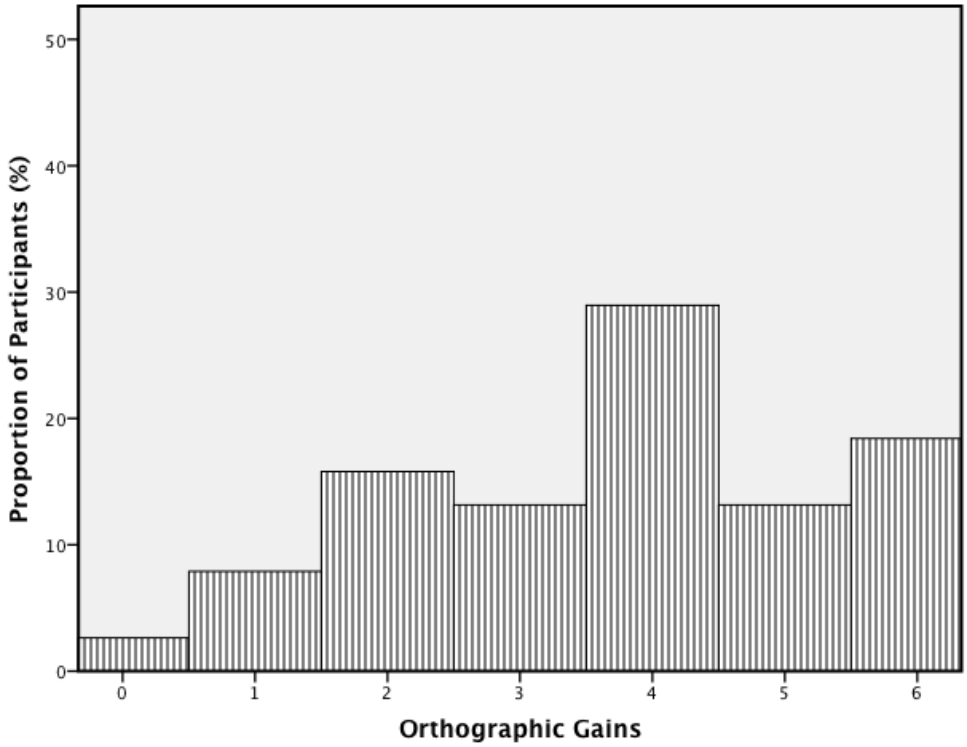

Figure 9: Distribution of orthographic gains in the video group

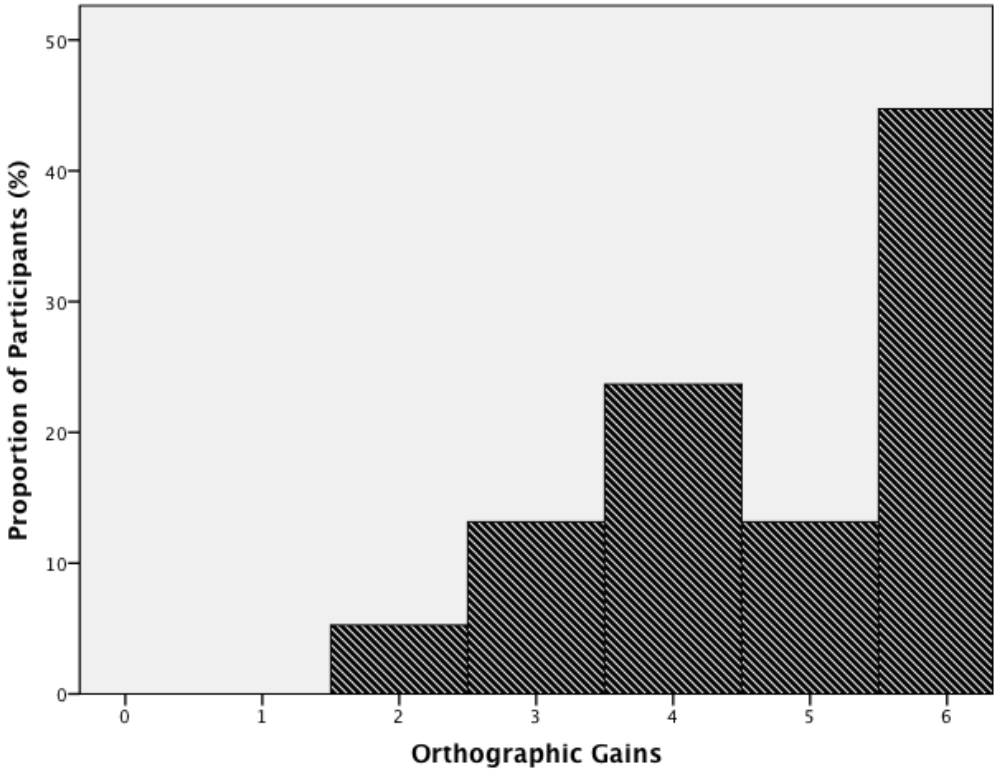

Figure 10: Distribution of orthographic gains in the blog group

4.2. Grammatical Function (Recognition)

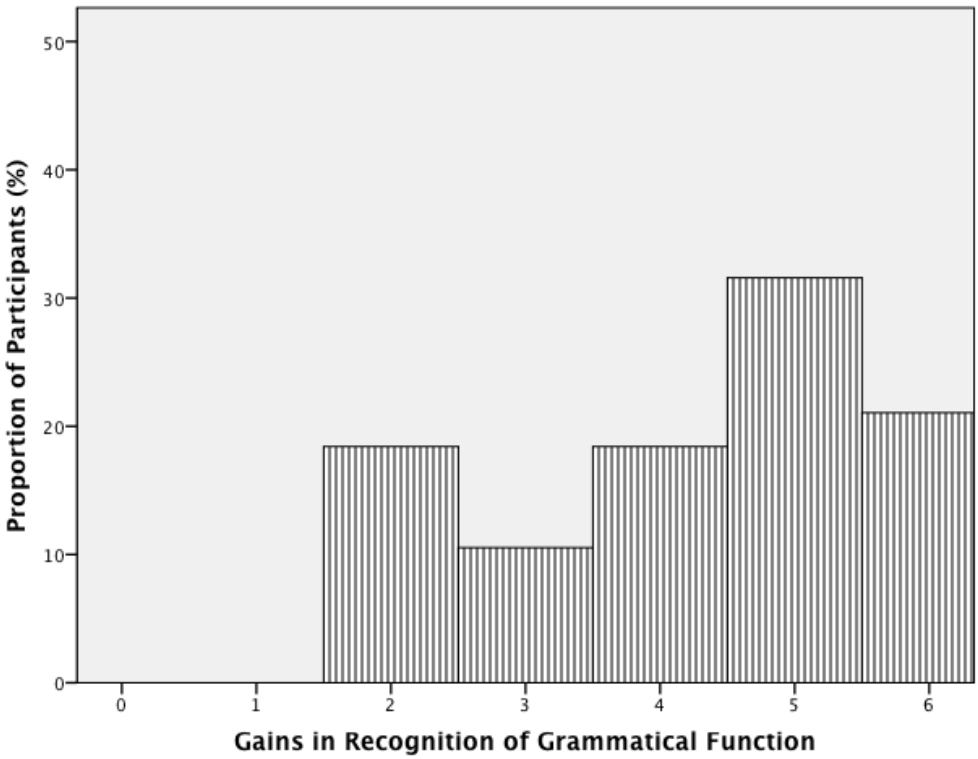

Figure 11: Distribution of gains in recognition of grammatical function in the video group

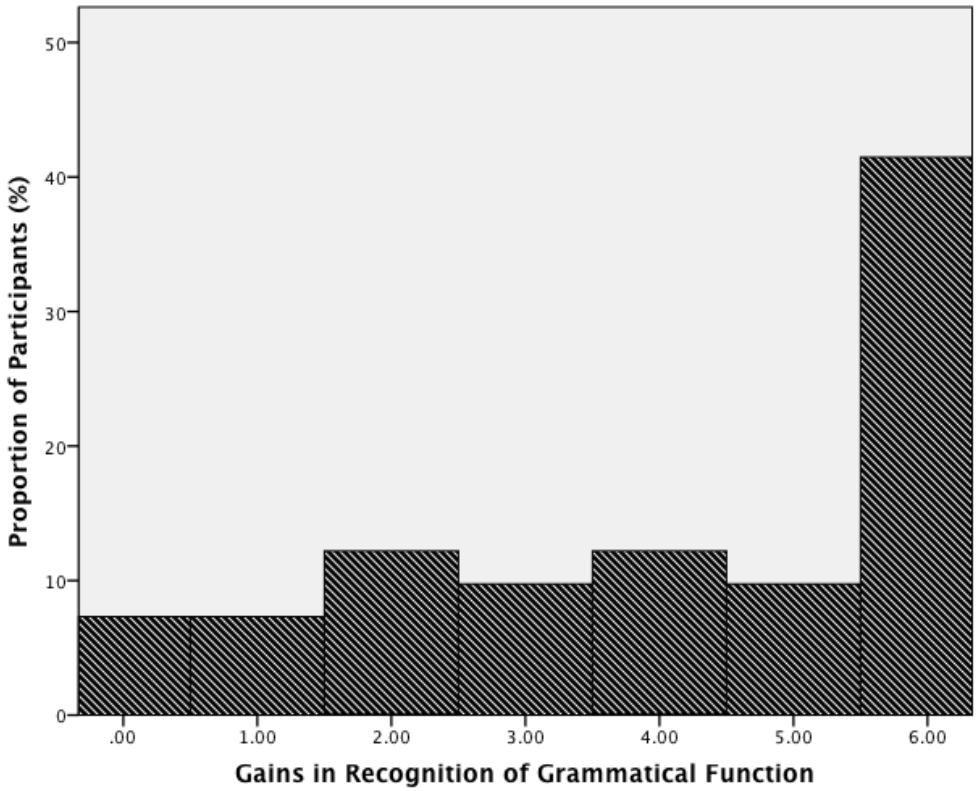

Figure 12: Distribution of gains in recognition of grammatical function in the blog group

4.3. Grammatical Function (Recall)

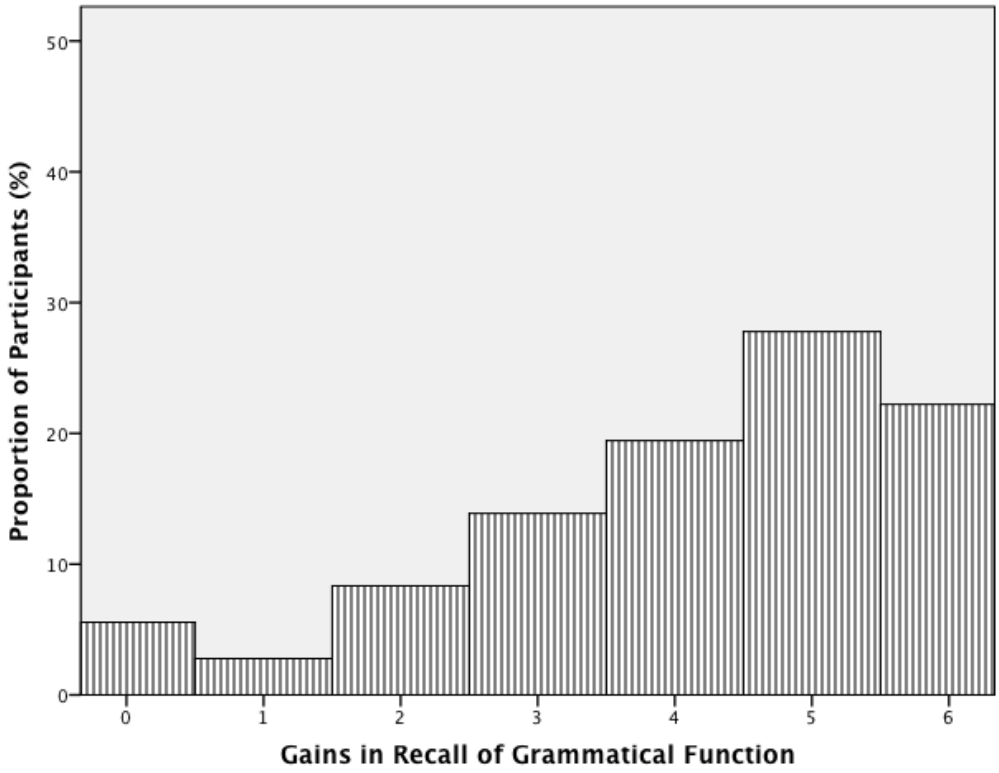

Figure 13: Distribution of gains in recall of grammatical function in the video group

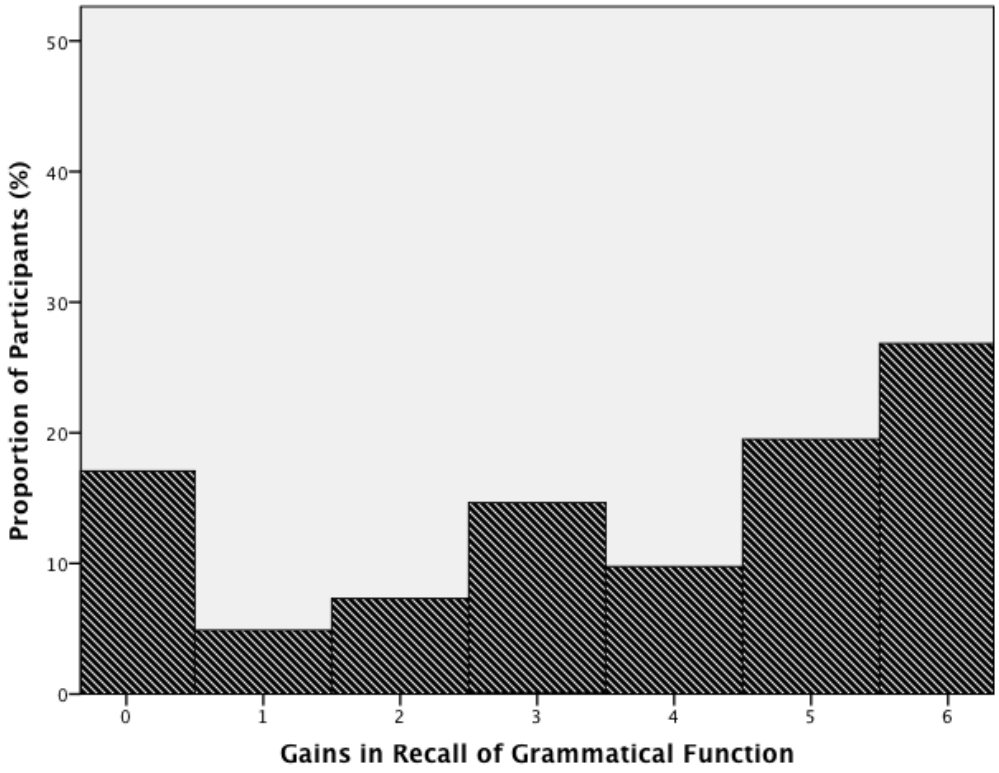

Figure 14: Distribution of gains in recall of grammatical function in the blog group

4.4. Meaning (Recognition)

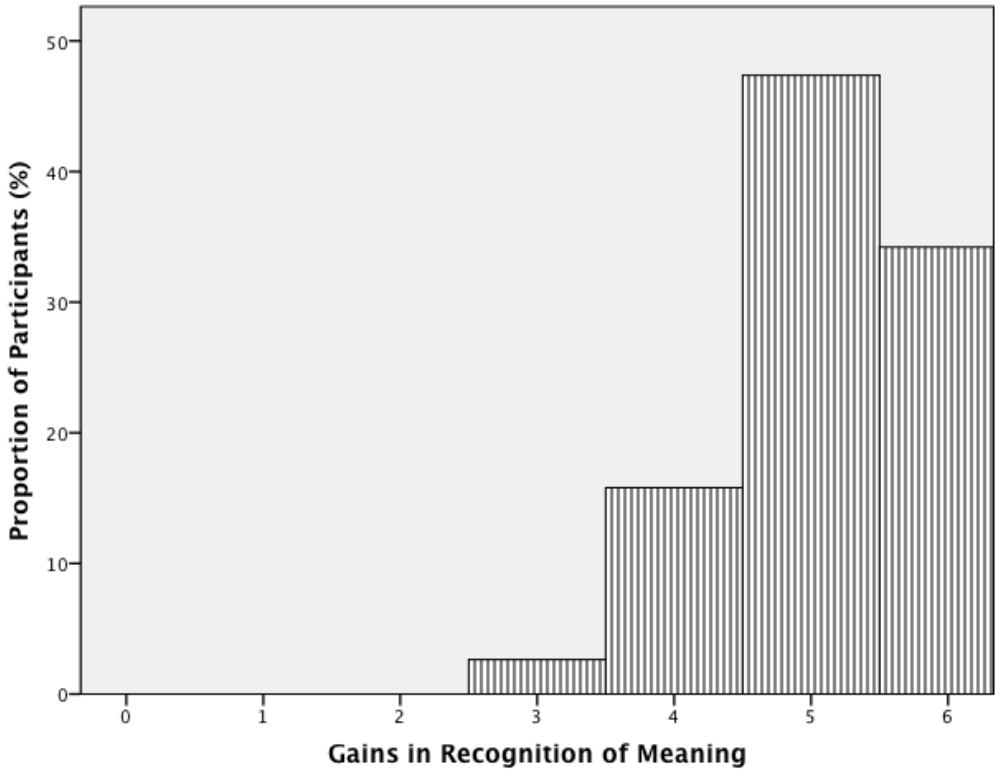

Figure 15: Distribution of gains in recognition of meaning in the video group

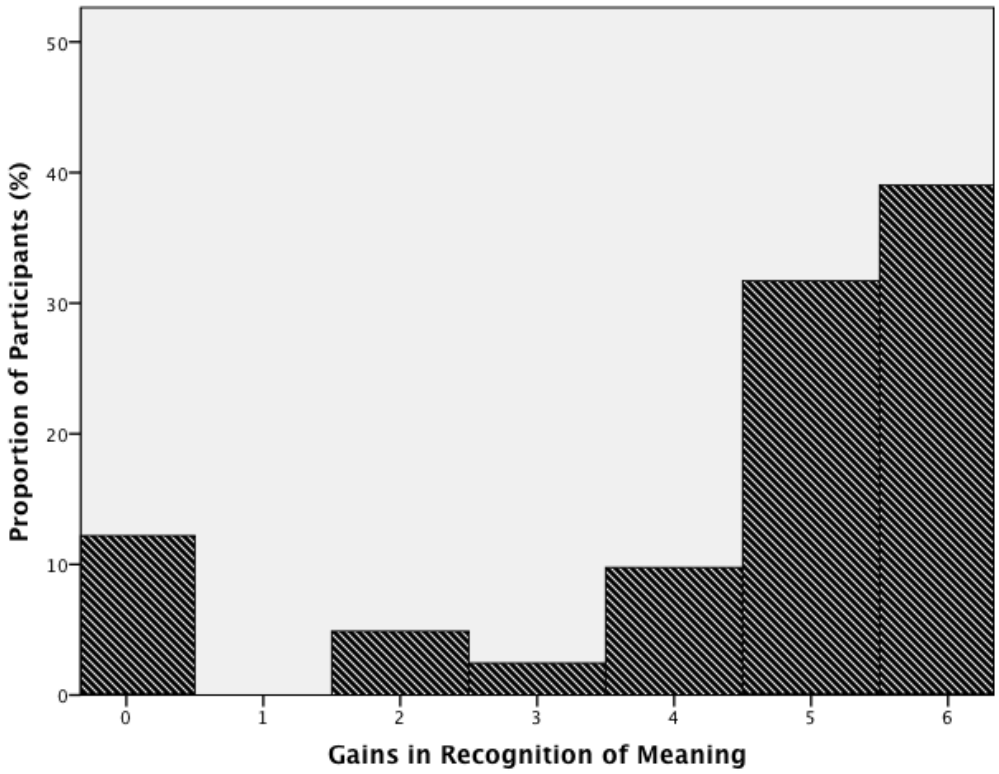

Figure 16: Distribution of gains in recognition of meaning in the blog group

4.5. Meaning (Recall)

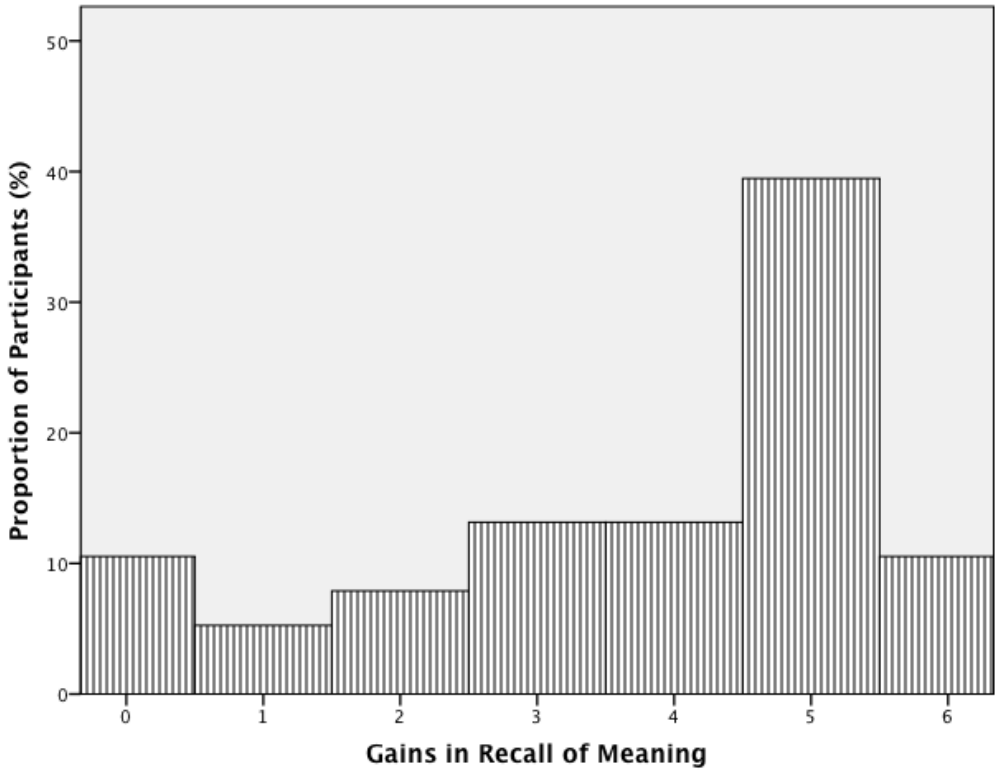

Figure 17: Distribution of gains in recall of meaning in the video group

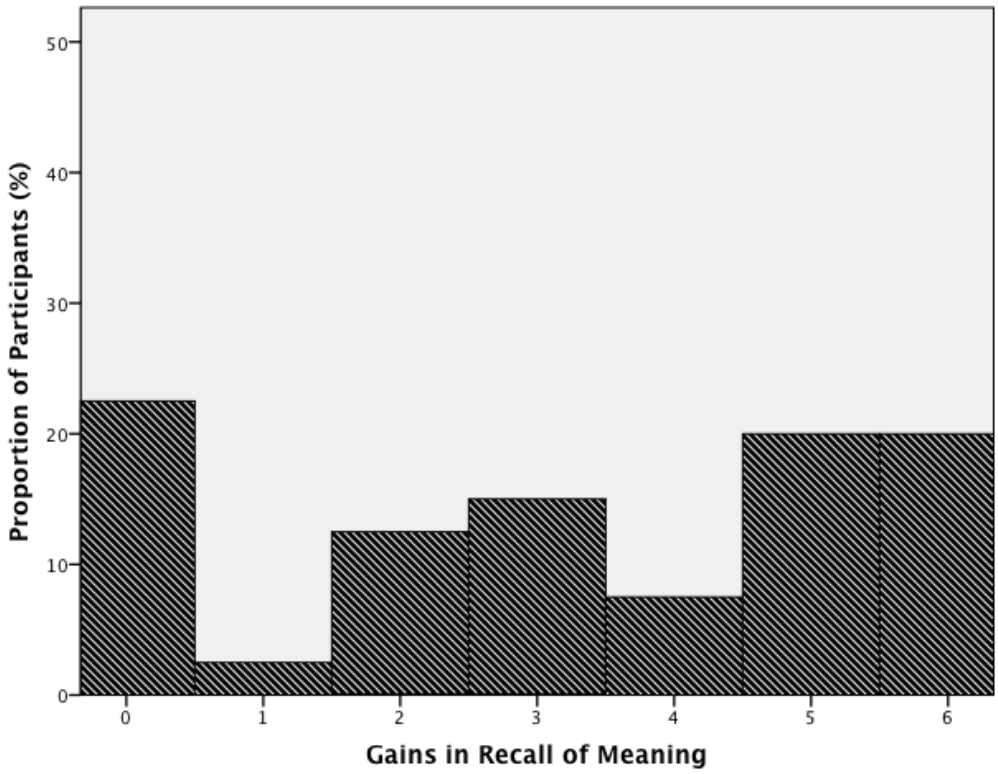

Figure 18: Distribution of gains in recall of meaning in the blog group