JAZYKY
V POHYBU
A POTENCIÁL
ZMĚNY

Sborník příspěvků z 19.–23. ročníku
mezinárodního setkání
mladých lingvistů

BOŽENA BEDNAŘÍKOVÁ
DAVID LAVIČKA
ZUZANA MALEŇÁKOVÁ
JIŘÍ MÁNEK
(eds.)

JAZYKY V POHYBU A POTENCIÁL ZMĚNY

Sborníky z konferencí
Filozofické fakulty, sv. 3

Filozofická
fakulta

Univerzita Palackého
v Olomouci

JAZYKY V POHYBU A POTENCIÁL ZMĚNY

Sborník příspěvků z 19.–23. ročníku mezinárodního setkání mladých lingvistů

BOŽENA BEDNAŘÍKOVÁ
DAVID LAVIČKA
ZUZANA MALEŇÁKOVÁ
JIŘÍ MÁNEK
(eds.)

Olomouc 2023

The Effects of Glossed and Unglossed Video Captions on L2 Vocabulary Learning

ZUZANA NADOVA

ABSTRACT

Previous research has demonstrated the benefits of using captions in instructional materials for learners' reading and listening comprehension, vocabulary acquisition and acquisition of grammar. However, studies comparing the effects of different captioning styles (audio, textual, L1, L2, keyword, full, glossed, unglossed captions) on vocabulary acquisition have yielded mixed results. Moreover, few studies have investigated the effectiveness of different captioning styles on vocabulary acquisition by using different types of recognition and production tasks.

The presented paper examines the results of a small-scale experiment studying the effectiveness of three captioning conditions created in English videos by means of the H5P authoring tool: unglossed captions with a highlighted target word, full captions with an L1 translation gloss of a target word and full captions with an L2 definition gloss of a target word. The experiment was conducted in an online instructional setting with Slovak EFL learners (N = 60) matched in age (m = 22.4) and proficiency measured by a standardized placement test score. The effects of the three captioning styles were measured in three immediate posttests aimed at form recognition, meaning recognition and contextual use of the target words. Statistical analyses (multivariate ANOVAs followed by univariate ANOVAs with pairwise comparisons) revealed (1) no statistically significant difference between the effects of the three types of captions on form recognition score; (2) a statistically significant effect of both types of glossed captions on meaning recognition score; (3) a statistically significant effect of glossed captions with L2 definition of a target word on contextual use score. The findings are discussed in the light of the Involvement Load Theory (Laufer & Hulstijn, 2001).

KEYWORDS

Vocabulary acquisition, video captions, glosses, EFL (English as a Foreign Language)

1. Introduction

In recent years, an important line of research in the field of second language (L2) and foreign language (FL) grammar and vocabulary acquisition in instructional settings has revolved around experimental investigation of the effects of using instructional technology in L2/FL learning[1] (Jung et al. 2017, p. 287). The significance of this line of research has increased during the academic year 2020 – 21, when most educational institutions had to adapt their teaching to online or hybrid teaching modes.

A frequently used method of applying instructional technology in L2 and FL instructional settings has been the use of captions and glosses[2] in textual and audio-visual materials. Vocabulary glosses, defined as "information provided about an unfamiliar linguistic item in the form of a definition, synonym, or translation, in order to reduce linguistic obscurity, and in doing so, assist reading comprehension" (Jung 2016, p. 93), may enhance multiple aspects of L2 and FL learning. Glosses draw learners' attention to both meaning and form of new vocabulary (Marzban 2011, p. 73), they enable the learner to match his own learning rate with finding new words (Marzban 2011, p. 76), they offer the opportunity for learners to engage in independent, incidental and continuous learning outside of class (Feng Teng 2020, p. 2), they reduce study time (Marzban 2011, p. 74), they are motivational in L2 reading because they make reading authentic texts more enjoyable (Ercetin 2003; Sakar & Ercetin 2005), etc. In addition, glosses have been shown to have a positive effect on many aspects of L2 and FL learning, including learners' reading comprehension (Marzban 2011; Salaberry 2001; Chen

---

1   In this paper, the term L2 learning/acquisition is used to refer to "a setting in which the language to be learnt is used by the local community" and FL learning/acquisition is used to refer to "a setting in which the language to be learnt is not the language spoken by the local community" (de Bot, Lowie & Verspoor 2005, p. 7).

2   In the current study, the term *caption* refers to on-screen textual information. Captions can be presented as onscreen textual captions in learners' L1 (translation captions) or L2 (definition captions). According to the Feng Teng (2020), L2 captions can lead to a triple asociation of sound, text, and image, thus enhancing vocabulary learning. The term *gloss* refers to an L1 translation or an L2 definition of a new vocabulary item embedded in a caption. A gloss can be visualised by clicking on the target word. In the present study, captions are presented in English and glosses are presented in textual (not audio) form either as English definition glosses of a target word or as Slovak translation glosses of a target word.

2016), vocabulary acquisition (Peters & Webb 2018, Chen 2016) and even grammatical competence (Jung 2016)[3]. The effects of audio and textual glosses have been studied both in textual teaching materials (Jung 2016; Rassaei 2017; Chen 2016), and in audio-visual materials (He 2019; Hsieh 2020; Hsu 2018; Feng Teng 2020; Montero Perez, Peters & Desmet 2018). However, studies comparing different captioning styles (glossed vs. unglossed captions, L1 vs. L2 captions, full vs. keyword captions, etc.) on vocabulary acquisition have produced inconclusive results (Hsieh 2020, p. 568; Jung 2016, p. 94). In addition, the relation between posttest task types and the effects of different types of glosses has not been sufficiently addressed by previous studies (Chen 2016, p. 416).

The current study aims to contribute to this field of research by comparing the effects of three types of video captions (unglossed captions with a highlighted target word, full captions with a first language (L1) translation gloss of a target word and full captions with an L2 definition gloss of a target word) on three aspects of vocabulary retention (form recognition, meaning recogniton and contextual use of the target words) measured by three immediate posttests. The study aims to identify differences between the three caption types in terms of their effectiveness for incidental vocabulary learning in English as a FL (EFL) and consequently to contribute to the field of knowledge about vocabulary acquisition in online learning environments with new data from Slovak EFL learners.

2. Theoretical Concepts

The effectiveness of different task types on vocabulary acquisition can be related to several key concepts. The Involvement Load Hypothesis proposed by Laufer and Hulstijn (2001) claims that retention of unfamiliar words is proportionate to the amount of *task-induced involvement load* while processing these words (Laufer & Hulstijn 2001, pp. 1–3). The authors operationalise *involvement* as a motivational-cognitive construct in terms of the degree of *need* to understand the lexical

---

3   On the other hand, other studies found no evidence of the effectiveness of glosses for the acquisition of grammatical structures. For instance, Martínez-Fernández (2010) found no effect of marginal L1 glosses on the learning of the Spanish subjunctive, however, they marginal L1 glosses were effective for the learning of vocabulary items. Similarily, Guidi (2009) found that glosses had only a limited effect on the learning of the Spanish *pretérito perfecto* and the impersonal *se.*

item in order to complete the task, *search* to infer the meaning of an unknown word and *evaluation* of a grammatical and semantic fit of the item (Laufer & Hulstijn 2001, p. 14). The rationale behind the Involvement Load Hypothesis can be explained by the premise that *elaboration* on features of new words promotes their retention (Anderson 1995; Baddeley 1997).

Another theoretical aspect of FL vocabulary learning is the difference between *incidental vocabulary learning* and *intentional vocabulary* learning. Incidental vocabulary learning occurs during a task, in which learners primarily attend to meaning, such as a reading comprehension task in which learners are focused on understanding the meaning of the text rather than committing unknown words to the memory while *intentional vocabulary learning* is defined as a deliberate word study task, such as a word list (Sok & Han 2020, p. 116). While intentional vocabulary learning enables learners to acquire explicit knowledge of form-meaning links of new words and gives rise to fluent access to these form meaning links (Sok & Han 2020, p. 117), incidental vocabulary learning enables learners to consider other information about the new words beyond form-meaning links, such as their semantic and syntactic relation to other words (Sok & Han 2020, p. 118).

In the current study, retention of new vocabulary from unglossed captions with a highlighted target word involves a higher degree of inferencing processes needed to derive the meaning of new words from the context while both types of glossed captions contain L1 translations and L2 definitions of the target vocabulary, respectively, thus reducing the need of inference from the context and providing learners with direct access to the meaning of a new lexical item.

Importantly, the three types of captions differ in terms of degree of task-induced involvement load along the dimension of evaluation. Specifically, both glossed conditions increase the evaluation of new lexical items since they allow learners to internalise their meaning by comparing them with their L1 translations or L2 definitions..

3. Previous Studies

Many empirical studies have addressed the relation between various gloss types and L2/FL learners' acquisition of vocabulary, grammar, and their reading comprehension skills. Chen (2016) investigated the effects of three gloss types (in-text, marginal and pop-up glosses) in texts on reading comprehension and vocabulary

acquisition in four task types (summary writing, multiple-choice questions, word translation and word matching). The results demonstrated that the same gloss mode yielded different outcomes in different task types. While marginal glosses led to the highest score in the multiple-choice test, in-text glosses led to the best score in the rest of the tests and pop-up glosses led to the lowest score in all four tests (Chen 2016, p. 419). A similar result has been obtained by Jung (2016) who found out that textual translation glosses had a positive effect on the acquisition of grammar (English unaccusative verbs) and vocabulary (10 pseudo-word items) but no effect on reading comprehension in Korean college students with an intermediate level of English. The study further revealed that glossing may make a differing contribution to the learning of form and meaning, although both aspects of vocabulary acquisition were enhanced by the glossing conditon.

The effectiveness of L1 vs. L2 glosses during reading comprehension has been investigated by various authors, showing mixed results. While some studies have shown that L1 glosses are more effective (Laufer & Shmueli 1997), other studies have indicated the opposite (Miyasako 2002) or even no difference between L1 and L2 glosses (Yoshii 2006). In terms of different types of post-tests, it has been shown that L1 captions might be more beneficial than keyword captions or no captions in terms of acquiring the contextual use of target words, for example in helping learners comprehend and reproduce orally the content of the video (He 2019).

Two recent large-scale studies investigating the effect of various captioning styles in videos are Montero Perez, Peters & Desmet (2018) and Hsieh (2020). The first study conducted with 277 students of Law or Economics enrolled at a Flemish University and learning EFL showed that keyword captioning and glossed keyword captioning led to a significant increase in posttest scores in form recognition. In addition, glossed captioning was most significant in both form and meaning recognition tests. The effectiveness of glossed captions for both short-term and long-term acquisition of Business English vocabulary has also been shown by Hsu (2018) in his study with Korean college learners of EFL. The second study conducted by Hsieh (2020) with 105 Chinese undergraduate EFL learners revealed that both full captions and L1 glossed full captions yielded similar results in form recognition task while in the meaning recognition and meaning recall tasks, L1 glossed full captions led to higher scores than full captions.

These findings indicate that various gloss types may significantly enhance different aspects of vocabulary acquisition and may lead to different performance

in different task types. The present study attempts to shed more light on the connection between various types of captioning and different aspects of vocabulary acquisition measured by three posttest tasks.

4. Method

4.1 DESIGN

The current study was designed to answer the following research questions (RQs):

RQ1. What type of underlying mechanism contributes to a better acquisition of new vocabulary items by EFL learners: inference of the meaning of a new lexical item from its context (unglossed captions with a highlighted target word) or glossing of new vocabulary items in learners' L1 or L2 (full captions with an L1 translation gloss of a target word vs. full captions with an L2 definition gloss of a target word)?

RQ2. Do different types of captioning lead to different results in the three posttest tasks focused on different aspects of vocabulary acquisition (form recognition, meaning recognition, contextual use)?

4.2 CONTEXT AND PARTICIPANTS

To answer these questions, 60 Slovak EFL learners were selected from among university students enrolled in various elective ESP courses at a Slovak university and majoring in Public Administration or Political Science. The participants were matched in proficiency measured by a Quick Oxford Placement Test at the beginning of the semester. The proficiency level of the participants selected for the experiment was intermediate B1 (vocabulary test score = 41 – 60 out of 120) and their age range was between 19 and 25 years (mean = 22.4).

The following vocabulary items were selected for the experiment: *assets, liabilities, determine, equity, suppliers, redundancy, plummet, soar, assemble, commodity, intangible, tariff.* The participants were not aware of the purpose of the study.

4.3 INSTRUMENTS

In total, 3 videos were presented to participants during one online ESP class. The videos entitled "Assets & Liabilities: Rich Dad, Poor Dad by Robert Kiyosaki" (3'02"), "The First Financial Bubble: The Tulipomania" (1'15") and "International Trade Explained"(6'42") were downloaded from Youtube and three versions of

each video were created by means of H5P authoring tool: a version with unglossed captions with a highlighted target word (an English caption with a target vocabulary item in *ITALIC CAPITAL* font), a version with full captions with an L1 translation gloss of a target word (an English caption with a Slovak translation gloss of a target vocabulary item), a version with full captions with an L2 definition gloss of a target word (an English caption with English definition gloss of a target vocabulary item). Examples of the three conditions can be seen in *Appendix 1*. All the captioned videos were presented with audio narration (bimodal input). The three versions of the videos differed only in their caption types, and they contained the same target vocabulary. Each participant was exposed to only one experimental condition and watched the three videos only once.

Since the effect of a didactic intervention should never be assessed by analyzing a single task (Chen 2016; Cheng & Good 2009; Sampedro Mella & Sánchez Gutiérrez 2016), three posttests were applied in the present study: 1. A form recognition task (a series of target vocabulary items from the videos with correct and incorrect spelling where participants had to indicate whether the presented items were real English words or not), 2. A meaning recognition task (a multiple-choice test where participants had to match the correct target word from the video with its definition), 3. A contextual use test (a fill-in-the gaps task where learners had to complete a text with the missing target words from the video listed above the text). All three groups took the three tests administered online as immediate timed posttests. These posttest tasks slightly differ in terms of their involvement load. Specifically, while *need* and *search* are constant across the tasks, a multiple-choice test and a fill-in-the-gaps test involve a higher degree of *evaluation* than the form recognition task since learners must compare and evaluate all the word meanings against the definition or against the context.

5. Procedure

The participants, randomly divided into three experimental conditions, individually watched the three videos during one online ESP class in April 2021. Participants in groups 2 and 3 (the two glossed conditions) were instructed to click on highlighted words in the caption to visualise and read the gloss. Group 1 watched videos with unglossed English captions containing a highlighted target word (in italic capital font). Group 2 watched videos with full captions with L1 translation

glosses and group 3 watched videos with full captions with L2 definition glosses. The participants were instructed to watch the videos without consulting any on-line dictionaries and they were not told that their knowledge of the vocabulary included in the videos would be tested afterwards. The three groups were exposed to the same target vocabulary.

Subsequently, the three groups took a test measuring three aspects of vocabulary retention: form recognition, meaning recognition and contextual use of the target words.

6. Results

Descriptive statistics for the performance of the three groups in the three posttest tasks are summarized in table 1 below:

Table 1. Descriptive statistics for the three groups and their performance in the three posttest tasks.

	GROUP 1: L2 unglossed captions (N = 20)	GROUP 2: L1 glossed captions (N = 20)	GROUP 3: L2 glossed captions (N = 20)
FORM	M = 6.05, SD = 1.39	M = 6.70, SD = 1.08	M = 6.55, SD = 1.27
MEANING	M = 4.85, SD = 1.13	M = 6.45, SD = 1.39	M = 6.95, SD = 1.19
CONTEXT	M = 3.95, SD = 1.31	M = 4.65, SD = 1.08	M = 6.20, SD = 1.73

Notes: GROUP 1 = unglossed captions with a highlighted target word, GROUP 2 = full captions with an L1 translation gloss, GROUP 3 = full captions with an L2 definition gloss. FORM = form recognition task, MEANING = meaning recognition task, CONTEXT = contextual use task. M = mean, SD = standard deviation.

The data in table 1 show that in the form recognition task, the highest mean score and the lowest standard deviation (M = 6.7, SD = 1.08) was achieved by group 2 (captions with L1 translation glosses), followed by group 3 (captions with L2 definition glosses, M = 6.55, SD = 1.27) and group 1 (unglossed captions with a highlighted keyword, M = 6.05, SD = 1.39). In the meaning recognition task, group 3 (M = 6.95, SD = 1.19) outperformed group 2 (M = 6.45, SD = 1.39) and group 1 (M = 4.85, SD = 1.13). The contextual use task scores show a similar

pattern as the meaning recognition task scores: the highest mean score can be attributed to group 3 (M = 6.2, SD = 1.73), followed by group 2 (M = 4.65, SD = 1.08) and group 1 (M = 3.95, SD = 1.31).

In summary, these results seem to indicate that both types of glossed captions lead to a better performance in all three posttests than unglossed captions. Unglossed captions seem to have a positive effect only on form recognition, at the expense of meaning recognition and contextual use.

In order to find out which differences between the three groups summarized in Table 1 were statistically significant, a MANOVA multivariate test was conducted in SPSS 20.0 (IBM, 2011) with three dependent variables (form, meaning, context) and one fixed factor (gloss type: unglossed captions with a highlighted target word, full captions with an L1 translation gloss of a target word and full caption with an L2 definition of a target word). Table 2 shows that the test yielded a statistically significant difference in the performance of the three groups in the three task types, $F (6, 110) = 7.006$, $p < 0.001$; Wilks $\Lambda = 0.523$, partial $\eta^2 = 0.277$.

Table 2. Results of one-way MANOVA.

Effect	Value	F	Hypothesis (df)	Error (df)	Sig.	Partial η^2
Intercept:						
Pillai's Trace	.977	776.180[b]	3.000	55.000	.000	.977
Wilk's Lambda	.023	776.180[b]	3.000	55.000	.000	.977
Hotelling's Trace	42.337	776.180[b]	3.000	55.000	.000	.977
Roy's Largest Root	42.337	776.180[b]	3.000	55.000	.000	.977
GROUP						
Pillai's Trace	.520	6.560	6.000	112.000	.000	.260
Wilk's Lambda	.523	7.006[b]	6.000	110.000	.000	.277
Hotelling's Trace	.827	7.445	6.000	108.000	.000	.293
Roy's Largest Root	.710	13.253[c]	3.000	56.000	.000	.415

Notes: F = test statistic, df = degrees of freedom, Sig = statistical significance. [b] exact statistic, [c] the statistic is an upper bound on F that yields a lower bound on lower significance level.
GROUP = caption type. Statistical significance: $p < 0.05$.

In order to obtain a more detailed account of the differences between the three groups in their performance in the three tasks, univariate ANOVAs were performed in SPSS. The results are shown in Table 3 below.

Table 3. Results of univariate ANOVAs (Tests of Between-Subjects Effects).

Source	Dependent Variable	df	F	Sig.	Partial η²
Corrected Model	FORM	2	1.466	.240	.049
	CONTEXT	2	13.411	.000	.320
	MEANING	2	15.509	.000	.352
Intercept	FORM	1	1570.990	.000	.965
	CONTEXT	1	738.555	.000	.928
	MEANING	1	1430.907	.000	.962
GROUP	FORM	2	1.466	.240	.049
	CONTEXT	2	13.411	.000	.320
	MEANING	2	15.509	.000	.352

Notes: GROUP = caption type (1 = unglossed captions with a highlighted target word, 2 = full captions with an L1 translation of a target word, 3 = full captions with an L2 definition of a target word). FORM = form recognition task, CONTEXT = contextual use task, MEANING = meaning recognition task. F = test statistic, df = degrees of freedom, Sig = statistical significance. Statistical significance after Bonferroni correction: $p < 0.017$.

Table 3 shows that group (caption type) yielded a statistically significant effect on the score in the contextual use task ($F_{(2,2)} = 13.411$; $p < 0.001$; partial $η^2 = 0.320$) and in the meaning recognition task ($F_{(2,2)} = 15.509$; $p < 0.001$; partial $η^2 = 0.352$). However, form recognition task score was not affected by group ($F_{(2,2)} = 1.466$; $p = 0.24$; partial $η^2 = 0.049$).

These significant ANOVAs were followed by Multiple comparisons with Tukey's HSD post-hoc tests detailed in table 4 below.

Results of the post-hoc analysis with pairwise comparisons shown in Table 4 indicate that there were no statistically significant differences between the groups in the form recognition task: group 1 vs group 2 (MD = -0.65; CI = -1.65, 0.35; p = 0.27), group 1 vs group 3 (MD = -0.5; CI = -1.5, 0.5; p = 0.46) and group 2 vs group 3 (MD = -0.15; CI = -0.85, 0.15; p = 0. 93).

Table 4. Multiple comparisons with Tukey's HSD post-hoc tests.

Dependent Variable	(I) GROUP	(J) GROUP	Mean Difference (I-J)	Sdt. Error	Sig	95% CI	
						Lower Bound	Upper Bound
FORM	1.00	2.00	-.6500	.39758	.239	-1.6067	.3067
		3.00	-.5000	.39758	.425	-1.4567	.4567
	2.00	1.00	-.6500	.39758	.239	-.3067	1.6067
		3.00	.1500	.39758	.925	-.8067	1.1067
	3.00	1.00	.5000	.39758	.425	-.4567	1.4567
		2.00	-.1500	.39758	.925	-1.1067	.8067
CONTEXT	1.00	2.00	-.7000	.44466	.265	-1.7700	.3700
		3.00	-2.2500*	.44466	.000	-3.3200	-1.1800
	2.00	1.00	.7000	.44466	.265	-.3700	1.7700
		3.00	-1.5500*	.44466	.000	-2.6200	-.4800
	3.00	1.00	2.2500*	.44466	.000	1.1800	3.3200
		2.00	1.5500*	.44466	.003	.4800	2.6200
MEANING	1.00	2.00	-1.6000*	.39392	.000	-2.547	-.6521
		3.00	-2.1000*	.39392	.000	-3.0479	-1.1521
	2.00	1.00	1.6000*	.39392	.000	.6521	2.5479
		3.00	-.5000	.39392	.418	-1.4479	.4479
	3.00	1.00	2.1000*	.39392	.000	1.1521	3.0479
		2.00	.5000	.39392	.418	-.4479	1.4479

Notes: FORM = form recognition task, CONTEXT = contextual use task, MEANING = meaning recognition task. GROUP 1 = unglossed captions with a highlighted target word, 2 = full captions with an L1 translation of a target word, 3 = full captions with an L2 definition of a target word. CI = Confidence Interval. * = p < 0.05.

On the other hand, meaning recognition task yielded statistically significant differences in mean scores between group 1 and group 2 (MD = -1.6; CI = -2.6, 0.6; p = 0.001) and between group 1 and group 3 (MD = -2.1, CI = -3.01; -1.1; p < 0.0001), but not between group 2 and group 3 (MD = -0.5; CI = -1.5; 0.5; p = 0.452).

In the contextual use task, mean scores were statistically significantly different between group 1 and group 3 (MD = -2.25; CI = -3.36; -1.13; p < 0.0001), and

between group 2 and group 3 (MD = -1.55; CI = -2.66, -0.43; p = .004), but not between group 1 and group 2 (MD = -0.7; CI = -1.81; 0.41, p = 0.297).

These results show that on the one hand, the three types of caption did not yield significantly different scores in the form recognition task but on the other hand, in the meaning recognition task, full captions with an L2 definition gloss of a target word contributed to a statistically significant increase in mean score in comparison with the unglossed condition. In the meaning recognition task, full captions with L1 translation glosses also led to a statistically significant increase in mean score in comparison with the unglossed condition, however, there was no statistically significant difference in mean scores between groups with captions with L1 translation glosses and captions with L2 definition glosses. Both glossed conditions significantly enhance meaning recognition scores in an immediate posttest task.

In addition, in the contextual use task, only full captions with an L2 definition gloss led to a significantly higher mean score in comparison with the other two conditions. Interestingly, there was no statistically significant difference in mean scores in the contextual use task between groups with captions with an L1 translation gloss and unglossed captions. Full glosses with an L2 definition gloss seem to be the most effective captioning style significantly enhancing contextual use of target words in an immediate posttest task.

7. Discussion and Conclusions

The results indicate that unglossed captions with a highlighted target word are the least effective caption type in terms of meaning recognition and contextual use of the target words, which implies that deriving the meaning of unknown words from the context seems to be a significantly less effective tool for incidental vocabulary learning from videos than glossing. This finding might be explained by the differences in the length of time spent by processing the three types of caption and the amount of attention devoted to the three captioning conditions. With both types of glossed captions, students had to click on the gloss and read it and so they were exposed to the target word and its definition or translation equivalent for a longer time than students in the unglossed condition.

The results have also shown that the effects of the three types of caption differ across the three posttest tasks. Form recognition task score was not significantly

affected by the three types of captions, however, meaning recognition and contextual use scores were significantly increased by both glossed caption types. This result is in line with the study by Hsieh (2020) who showed that full captions with a highlighted target word and an L1 translation gloss facilitated both the learning of form and meaning while full captions with highlighted target word increased only form recognition (Hsieh 2020, p. 567).

In the present study, full captions with an L2 definition gloss of a target word were shown to be the most effective captioning strategy for meaning recognition as well as for contextual use of the target words. A similar result was obtained by Montero Perez, Peters & Desmet (2018) who found out that the glossed keyword captions condition led to the most significant increase in form recognition and meaning recall tests (Montero Perez, Peters & Desmet 2018, p. 1).

In conclusion, the present study has shown that EFL learners benefit the most from glossed captions which significantly increase their posttest scores in meaning recognition and contextual use tasks in comparison with unglossed captions. Interestingly, while both types of glossed captions (with a higher degree of involvement load than unglossed captions) led to a similar posttest score increase in meaning recognition, L2 definition glosses (providing learners with direct access to vocabulary meaning in L2) were the most effective in increasing posttest scores in contextual use of new vocabulary items. The question whether tasks with higher involvement load such as the meaning recognition task and the contextual use task used in this study as posttests also affect short-term and/or long-term retention of the target vocabulary needs further empirical research. However, it may be hypothesized that using videos for incidental vocabulary learning with glossed captions combined with a posttest task with a higher degree of task-induced involvement might represent an effective method of vocabulary retention after a single intensive exposure and could also lead to long-term vocabulary retention. The task-induced involvement might be increased by introducing higher degress of *need* (the necessity to understand a target word in order to complete a task), *search* (e.g. looking up new vocabulary items or discussing their meaning with peers) and *evaluation* (comparing the lexical and semantic fit of an item in a specific context).

The present study has several limitations. One of them is a lack of data on multiple learner-related variables. For instance, besides measuring proficiency level, it would be useful to investigate possible correlations between the effects of glosses on the posttest task scores and other learner-related variables, such

as learners' frequency and amount of exposure to English, their motivation, and attitudes towards EFL, their visual working memory, and possibly others. A further limitation of the present study concerns the lack of comparison of the effects of glosses with a significantly different form of vocabulary learning, for example flashcards or a list of vocabulary with definitions. A possible extension of the present study could compare the results of the present experiment with an experiment conducted with EFL learners with different proficiency levels and/or from different age groups. And finally, a delayed posttest could be incorporated into the study design to test long-term retention of the target words across the three task types.

References

Anderson, J.R. (1995). *Cognitive Psychology and its Implications. 4ᵗʰ Edition*. New York: Freeman.

Baddeley, A. (1997). *Human Memory: Theory and Practice.* Hove: Psychology Press.

Chen, I.-Y. (2016). Hypertext Glosses for Foreign Language Reading Comprehension and Vocabulary Acquisition. Effects of Assessment Methods. *Computer Assisted Language Learning* 29(2): pp. 413–426.

Cheng, Y.H., & Good, R.L. (2009). L1 Glosses: Effects on EFL Learners' Reading Comprehension and Vocabulary Retention. *Reading in a Foreign Language* 21(2), pp. 119–142.

De Boot, K., Lowie, W. & Verspoor, M. (2005). *Second Language Acquisition. An Advanced Resource Book*. New York: Routledge.

Ercetin, G. (2003). Exploring ESL Learners' Use of Hypermedia Reading Glosses. *CALICO Journal*, 20(2), 261–283.

Feng Teng, M. (2020). Vocabulary Learning Through Videos: Captions, Advance-Organizer Strategy, and their Combination. *Computer Assisted Language Learning* 2020. DOI 10.1080/09588221.2020.1720253

Guidi, C. (2009). *Glossing for Meaning and Glossing for Form. A Computerized Study of the Effects of Glossing and Type of Linguistic Item on Reading Comprehension, Noticing and L2 Learning*. PhD dissertation, Georgetown University. Available online at:https://repository.library. georgetown.edu/handle/10822/553239

He, I. (2019). How Do Different Keyword Captioning Strategies Impact Students' Performance in Oral and Written Production Tasks? *The EuroCALL Review* 27(1): pp. 14–23.

Hsieh, Y. (2020). Effects of Video Captioning on EFL Vocabulary Learning and Listening Comprehension. *Computer Assisted Language Learning* 33 (5-6): pp. 567–589.

Hsu, H-T. (2018). Incidental Professional Vocabulary Acquisition of EFL Business Learners: Effects of Captioned Video with Glosses. *The JALT CALL Journal 2018* 14 (2): pp. 119–142.

Jung, J. (2016). Effects of Glosses on Learning of L2 Grammar and Vocabulary. *Language Teaching Reserch* 20(1): pp. 92–112.

Jung, J., Y. Kim, H. Lee, R. Cathey, J. Carver & S. Skalicky S. (2017). Learner Perception of Multimodal Synchronous Computer-Mediated Communication in Foreign Language Classrooms. *Language Teaching Research* 22 (3): pp. 287–309.

Laufer, B. & J. Hulstijn (2001). Incidental Vocabulary Acquisition in a Second Language: The construct of Task-Induced Involvement. *Applied Linguistics* 22(1): pp. 1–26.

Laufer, B., & Shmueli, K. (1997). Memorizing New Words: Does Teaching Have Anything to Do with It? *RELC Journal*, 28: pp. 89–108.

Martínez-Fernández, A.M. (2010). *Experiences of Remembering and Knowing in SLA, L2 Development and Text Comprehension: A Study of Levels of Awareness, Glossing and Type of Linguistic Item.* PhD dissertation, Georgetown University. Available online at: https://repository.library.georgetown.edu/bitstream/handle/10822/558184/MartinezFernandez_georgetown_0076D_10882.pdf;sequence=1

Marzban, A. (2011). Investigating the Role of Multimedia Annotations in EFL Reading Comprehension. *Procedia – Social and Behavioral Sciences* 28: pp. 72–77.

Miyasako, N. (2002). Does Text-Glossing Have any Effects on Incidental Vocabulary Learning through Reading for Japanese Senior High School Students? *Language Education & Technology*, 39: pp. 1–20.

Montero Perez, M., Peters, E. & P. Desmet. (2018). Vocabulary Learning through Viewing Video: The Effect of Two Enhancement Techniques. *Computer Assisted Language Learning* 31(1-2), pp. 1–28.

Peters, E. & S. Webb (2018). Incidental Vocabulary Acquisition through Viewing L2 Television and Factors that Affect Learning. *Studies in Second Language Acquisition* 40(3): pp. 551–577.

Rassaei, E. (2017). Computer-Mediated Textual and Audio Glosses, Perceptual Style and L2 Vocabulary Learning. *Language Teaching Research* 22(6): pp. 657–675.

Sampedro Mella, M.C. & C. H. Sánchez Gutiérrez (2016). Efecto de la enseñanza extensiva en la memorización del léxico: un estudio de investigación – acción. *Sintagma* 29: pp. 61–78.

Sakar, A., & G. Ercetin (2005). Effectiveness of Hypermedia Annotations for Foreign Language Reading. *Journal of Computer Assisted Learning*, 21(1), 28–38.

Salaberry, M.R. (2001). The Use of Technology for Second Language Learning and Teaching. A Retrospective. *The Modern Language Journal* 85(1): pp. 39–56.

Sok, S., & Z. H. Han (2020). A Study of L2 Vocabulary Acquisition Under Incidental and Intentional Conditions. *Vigo International Journal of Applied Linguistics* 17: pp. 113–140.

Yoshii, M. (2006). L1 and L2 Glosses. Their Effects on Incidental Vocabulary Learning. *Language Learning & Technology* 10: pp. 85–101.

Appendix 1. Examples of presentation of video captions: A. unglossed captions with a highlighted target word, B. full captions with an L1 translation gloss of a target word, C. full captions with an L2 definition gloss of a target word.
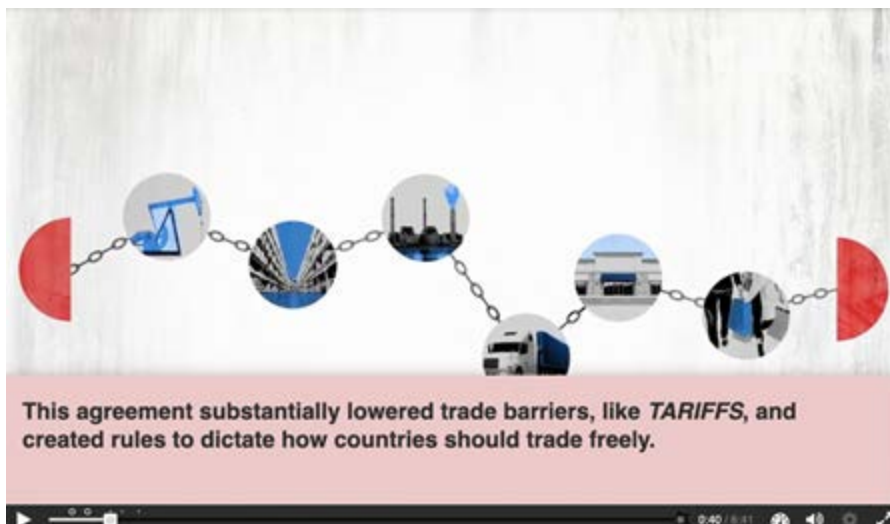
A. unglossed captions with a highlighted target word

B. full captions with an L1 translation gloss of a target word

C. full captions with an L2 definition gloss of a target word

VLIV GLOSOVANÝCH A NEGLOSOVANÝCH VIDEOTITULKŮ NA BUDOVÁNÍ SLOVNÍ ZÁSOBY V JAZYCE L2

Predchádzajúci výskum poukázal na efektívnosť využívania titulkov v didaktických materiáloch, vrátane ich pozitívneho vplyvu na schopnosti čítania a posluchu s porozumením, a na osvojenie si novej slovnej zásoby a gramatických štruktúr. Štúdie, ktorých cieľom bolo porovnanie efektívnosti rôznych typov titulkov pri osvojovaní si slovnej zásoby (napr. textové titulky, titulky vo forme audio súboru, titulky v materinskom jazyku, titulky v cieľovom jazyku, titulky vo forme kľúčového slova, titulky vo forme kompletného prepisu videa, titulky s možnosťou zobrazenia prekladu alebo definície nového slova, atď) často dospeli k rozporuplným výsledkom. Navyše, veľmi málo štúdií porovnávalo efektívnosť rôznych druhov titulkov prostredníctvom viacerých typov úloh, zameraných na viacero aspektov porozumenia a použitia nových slov.

Cieľom tohto príspevku je prezentovať výsledky experimentu, zameraného na porovnanie efektívnosti troch typov titulkov, vytvorených v anglických videách prostredníctvom nástroja *H5P authoring tool*: titulky vo forme kompletného prepisu videa v anglickom jazyku s graficky zvýrazneným kľúčovým slovom, titulky v anglickom jazyku vo forme kompletného prepisu videa s možnosťou zobrazenia anglickej definície nového slova a titulky v anglickom jazyku vo forme kompletného prepisu videa s možnosťou zobrazenia prekladu nového slova do slovenského jazyka. Účastníkmi experimentu boli slovenskí univerzitní študenti angličtiny ako cudzieho jazyka (počet = 60,

priemerný vek = 22.4 rokov) s mierne pokročilou úrovňou angličtiny, meranou štandardizovaným zaraďovacím testom. Efektívnosť titulkov bola meraná v troch úlohách, zameraných na identifikáciu správnej formy, identifikáciu významu a na správne použitie nových slov v kontexte. Štatistická analýza (multivariačná analýza rozptylu ANOVA s následnými univariačnými analýzami rozptylu ANOVA s plánovanými párovými porovnaniami) ukázala (1) žiaden štatistický významný rozdiel medzi vplyvom skúmaných tipov titulkov na identifikáciu správnej formy nového slova; (2) štatisticky významný efekt titulkov s možnosťou zobrazenia prekladu nového slova alebo jeho definície na skóre v úlohách, zameraných na identifikáciu správneho významu nových slov; (3) štatisticky významný efekt titulkov s možnosťou zobrazenia deficície kľúčového slova na skóre v úlohách, zameraných na správne použitie nových slov v kontexte. Výsledky experimentu sú vysvetlené v kontexte teórie *Involvement Load Theory* (Laufer & Hulstijn 2001).

KĽÚČOVÉ SLOVÁ
Osvojovanie si slovnej zásoby, video titulky, preklad a definícia kľúčového slova vo video titulkoch, anglický jazyk ako cudzí jazyk

KONTAKT NA AUTORA
Zuzana Nadova, M.A.
Faculty of Arts,
Pavol Jozef Šafárik University in Košice, Slovakia
zuzana.nadova1@upjs.sk