

Vocabulary learning from audiovisual input at first exposure in young adult novice learners

Imma Miralpeix, Ferran Gesa & Maria-del-Mar Suárez
Universitat de Barcelona

In this ‘First Exposure’ (FE) study, 106 Catalan/Spanish young adults proficient in English watched a short advert with the audio in English and subtitles in Polish, a language they were not familiar with. Results indicated that vocabulary learning took place, as their meaning recognition scores were significantly higher than those of a control group who had not seen the video. The most recurrent learning strategies these novice learners used were associations with imagery and with vocabulary in previously learned languages, along with attention to specific input factors such as frequency of word occurrence. Findings reveal how audiovisual input can help vocabulary acquisition at FE, as well as factors relevant for learning from video viewing at these very first stages.

1. Introduction

Recent advancements in digital technology have transformed the way we engage with new languages, and the prevalence of exposure to audiovisual materials has been steadily increasing in contemporary times (De Wilde et al., 2020). A growing body of research indicates that audiovisual input enhances second language (L2) learning (see, for instance, Reynolds et al., 2022). However, much of the existing research has predominantly focused on *en route* learners – those already in the process of learning a language. There is a dearth of research investigating whether and how audiovisual input facilitates learning for *ab initio* or novice learners – individuals encountering a language for the first time (see, for instance, Rast, 2008). We denote this area of study as ‘First Exposure’ (FE) research.

Conducting studies on FE holds relevance both theoretically and practically. Firstly, as highlighted by scholars like Gullberg and Indefrey (2010) or Han and Rast (2014), the initial phases of adult L2 vocabulary acquisition have often been overlooked in research, thus creating a gap in the literature investigating the nat-

ural processes and mechanisms underlying learners' capacity to 'break into the wild' (see, for instance, Reynolds, 2023). Furthermore, it is crucial to uncover the pre-linguistic knowledge individuals bring to the task of language acquisition, along with exploring input and learner factors that may facilitate learning during these early stages. From a practical standpoint, acquiring a core vocabulary early on is essential for effective communication (de Groot & van Hell, 2005), particularly when starting from scratch in a new language. Thus, any input source that aids in accelerating our learning rate is worth exploring, as noted by McLaughlin et al. (2004), who observed that minimal instruction can lead to rapid progress.

To date, FE studies have examined language acquisition using various types of input, primarily artificial languages (i.e., languages intentionally developed for experimental purposes, often in laboratory settings) or natural languages (typically in classroom settings). Examples of the former include studies by Cunillera et al. (2006), de Diego Balaguer et al. (2007), and Folia et al. (2010), while an instance of the latter is the VILLA project, where controlled classroom input was provided to French learners of Polish to evaluate their linguistic progress at multiple testing intervals (e.g., Dimroth et al., 2013; Rast, 2008, 2010a, 2010b; Shoemaker & Rast, 2013).

These FE investigations have revealed that lexical knowledge can indeed be acquired upon initial exposure to a completely unfamiliar language. For instance, Cunillera et al. (2006) demonstrated that both stress and statistical cues aid in word segmentation, while Shoemaker and Rast (2013) found that cognate words are recognized more readily than non-cognates, and that word recognition at FE is influenced by utterance position. However, there remains a significant gap in understanding how audiovisual input, consistently shown to promote language acquisition (e.g., Vanderplank, 2016), might contribute to intentional vocabulary learning during FE. Likewise, exploration is needed to uncover the strategies learners employ in this endeavour and their perceptions of the viewing experience.

2. Literature review

2.1 FE studies and audiovisual input

Two studies have thus far explored learning from audiovisual input during initial exposure, both focusing on incidental learning: Gullberg et al. (2010, 2012) and Bisson et al. (2014). In the former study, L1 Dutch participants viewed a seven-minute Chinese weather forecast (without on-screen text) specifically designed for the experiment, with several lexical items in the input enhanced by gestures. Results revealed participants' ability to extract both formal and semantic infor-

mation from limited input, as evidenced by word form and meaning recognition tests. The impact of word frequency and syllable structure on performance was also observed. Moreover, the use of fMRIs on participants during task performance yielded neurocognitive insights: for example, the involvement of the supramarginal gyri – regions within the inferior parietal lobule implicated in phonological processing – was noted in the creation of phonological representations of potential L2 words during brief exposure, accompanied by rapid neurological adaptation.

In Bisson et al. (2014), L1 English participants viewed a 25-minute episode of *SpongeBob* in Dutch under three distinct conditions: with intralingual subtitles (Dutch audio and Dutch subtitles), with standard subtitles (Dutch audio and English subtitles), and in a reversed subtitles condition (English audio and Dutch subtitles). Additionally, a control group watched the episode with Dutch audio and no subtitles. Using eye-tracking technology, the study confirmed the processing of subtitles and images, with participants dedicating less time to reading when the audio was in English. Nonetheless, no evidence of lexical learning was found in a vocabulary auditory test, where participants were presented with 78 Dutch words and had to determine whether the translation appearing on the screen for each word was correct or not.

Thus, in incidental learning conditions, Gullberg et al. (2010, 2012) demonstrated participants' ability to extract lexical knowledge from a brief, subtitled video, even if the languages under examination in the study were typologically distant, stemming from different language families (L1 Dutch – L2 Chinese). However, in contrast, Bisson et al. (2014) found no incidental vocabulary acquisition when participants viewed a lengthier video, despite the languages being typologically closer (L1 English – L2 Dutch).

It is worth mentioning that Hofweber et al. (2022, 2023) have also conducted research on incidental learning from audiovisual input at FE, focusing on the acquisition of sign language. These studies yielded interesting findings, such as the impact of input properties (e.g., item frequency or degree of iconicity) on language learning abilities during FE, surpassing the influence of other individual variables examined (e.g., intelligence or aptitude). However, given the distinct modalities of expression between sign and verbal languages, we do not provide an exhaustive account of these studies within this [paper](#).

2.2 Factors influencing learning at FE

2.2.1 *Input and learner factors*

Several factors have been shown to influence learning in FE studies, namely input- and learner-related factors. In relation to the first ones, we can mention syllable structure or gestural cues. For instance, participants in Gullberg et al. (2012) learned disyllabic words with stress on the first syllable more easily, as they were also more usually found in their L1 than monosyllabic words. Moreover, those words highlighted with gestures were also more liable to be learned. Utterance position has also been found to influence learning: as Shoemaker and Rast (2013) have indicated, novice learners in language classes are particularly attentive to the edges of prosodic domains. In contrast, word frequency has been shown to influence learning in Gullberg et al. (2012) and in Miralpeix et al. (2023), but not in Shoemaker and Rast (2013).

In terms of learner-related factors, previous studies have primarily focused on age, aptitude, languages previously learned, and psychotypology. For instance, Ristin-Kauffmann and Gullberg (2014) categorized their participants into nine age groups ranging from 10 to 86 years old. They discovered that performance in a lexical decision task, following exposure to audiovisual input, improved with age. Miralpeix et al. (2023), in a study involving a subset of participants from the current study's sample, found that linguistic aptitude (measured by the LLAMA B test; Meara, 2005) explained 24% of the variance in meaning recognition vocabulary scores. Moreover, Carroll (2014) noted that cognate stimuli were easier to segment and learn than non-cognate items. This advantage with transparent words has also been observed by Rast et al. (2014), suggesting that language typology and psychotypology (i.e., the perceived distance between languages; Kellerman, 1979) play a role in learning from minimal input (Rast, 2010b). It should be noted that, in research on learning from audiovisual input, learners' previous exposure to this type of input or their perceptions on the task performed have rarely been considered.

2.2.2 *Vocabulary learning strategies*

While prior research has begun to explore input and learner variables that could influence learning within foreign language (FL) education, learning strategies have received relatively little attention. Griffiths (2008) defined strategies as those activities consciously chosen by learners to regulate their own learning. Even if different researchers have classified strategies in different ways – see, for instance, Oxford (1990) – a difference is often made between cognitive strategies (operations on the incoming input to enhance learning, such as chunking, visualization or association), metacognitive (involving learners' awareness and control over

their own learning process, such as planning or evaluating), and affective (which involve managing emotions, attitude and motivation to facilitate learning, such as setting goals or building confidence).

Despite the abundance of studies on vocabulary learning strategies (e.g., Brown & Payne, 1994; Gu & Johnson, 1996; Pavičić Takač, 2008; Schmitt, 1997, 2000; Zhang & Li, 2011), fewer have focused on novice learners. For instance, in a study where English native speakers were tasked with reading while listening to a text in Dutch (a language they did not know), Singleton and Little (1984) observed that learners relied on their prior linguistic knowledge and input features simultaneously, and that they especially fell back on those previously learned languages typologically closer to the new one. In relation to the strategies used when exposed to multimodal input, the studies have been conducted with *en route* learners. For example, Borrás and Lafayette (1994), assessing the effects of subtitles on learner's production, posited that a low anxiety environment, as well as students' familiarity with this type of input, can contribute to the use of affective strategies. Furthermore, Thompson and Rubin (1996) claimed that learners engaging with audiovisual input can employ both cognitive (e.g., using visual cues for predictions or identifying cognates) and metacognitive strategies (e.g., setting listening objectives, planning and establishing goals). Sydorenko (2010) found out that beginner learners were acquiring most words by associating captions with images. In extensive viewing, where learners regularly watched original version TV, Vanderplank (2016) noticed that those who made use of retention strategies (e.g., keeping record of new words) learned more than those who did not. So far, then, we lack information on what novice learners do to try to understand messages conveyed by subtitled audiovisual input and to retain the information that they may gather from it.

2.2.3 *Learners' perception of the viewing experience*


While some studies have explored learners' perceptions of the effectiveness of L2 audiovisual materials, research in FE situations is lacking. Studies involving *en route* learners have commonly found these materials to be beneficial for learning, particularly in terms of vocabulary acquisition (e.g., Kusy & Sockett, 2012; Stewart & Pertusa, 2004). Additionally, it has been noted that learners may not always recognize that learning is occurring (e.g., Sydorenko, 2010, in the context of learning from short videos) or be aware of which viewing condition fosters better learning outcomes (e.g., Pattemore et al., 2024, comparing captions, enhanced captions, and no captions).

Despite the lack of research, the study of learners' perceptions is particularly pertinent in learning from viewing at FE. The initial exposure to a new language can be daunting for novice learners, potentially affecting their engagement with

the learning material, confidence levels, self-efficacy beliefs regarding language learning, and, consequently, their overall performance.

3. Research questions

Given the identified gaps in the literature and the abundance of audiovisual materials now available to learners in multiple languages, this chapter aims to explore the potential benefits of this type of input for novice learners. For this purpose, we pose the following research questions:

- RQ1. Can adult Catalan/Spanish native speakers establish word form-meaning connections after minimal exposure to audiovisual input with the audio in a language they know (English) and subtitles in a completely new language (Polish)?
- RQ2. In relation to vocabulary learning in the novel language 
 - RQ2a. Is it related to knowledge of other FLs learners may know or to previous exposure to audiovisual input in other languages?
 - RQ2b. Which vocabulary learning strategies do learners use?
- RQ3. How do novice learners perceive this viewing/learning experience?
 - RQ3a. Does learning depend on learners' perceptions?

4. Method

4.1 Participants


135 Catalan/Spanish pre-service teachers pursuing studies in early childhood and primary education at the University of Barcelona participated in the study: 106 belonged to the Experimental Group (EG) and 29 to the Control Group (CG); see Table 1. At the time of data collection, they had studied English at primary and secondary school for about ten years and had a mean receptive vocabulary size of about 4,500 words, according to  V_YesNo test (Meara & Miralpeix, 2015). Moreover, there were not any significant differences between the two groups as far as their English vocabulary size was concerned, as corroborated by an independent samples *t*-test (see Table 1). They did not have any knowledge of Polish and had never been exposed to it either formally or informally, as was confirmed by a questionnaire that participants filled in after taking the tests (Appendix B) and by direct questions from the researchers to the participants in case of any doubt.

Table 1. Participants in the study

	<i>N</i>	Age <i>M (SD)</i>	<i>L1</i>	English vocabulary size <i>M (SD)</i>	
Experimental (EG)	106	19 (.667)	Catalan/Spanish	4,554 words (1,410.29)	$t(71.822) = .488,$ $p = .627$
Control (CG)	29	18.7 (.556)	Catalan/Spanish	4,443 words (931.76)	

4.2 Instruments

4.2.1 TV advert

A TV advertisement from the renowned e-commerce platform Allegro, accessible at <https://rb.gy/9phftk> (Allegro, 2023), served as the stimulus for this study. The advert, lasting 2 minutes and 58 seconds, depicted the story of a resourceful Polish grandfather preparing for a trip to the UK, and learning English at home to be able to communicate with his granddaughter. The audio of the advert was in English, while the subtitles were provided in Polish. The subtitled text comprised 75 tokens and 45 types (i.e., distinct words) across a total of 34 subtitles, ranging from one to six words in length, and with a mean of 2.2 words per subtitle. In terms of the audio content, approximately 95% of the words fell within the 1k frequency band, with the remaining 5% in the 2k band, according to VocabProfile (Cobb, n.d.). This distribution aligned well with the learners’ competence level, as their receptive vocabulary size exceeded by far 2,000 words.

It is important to note that the technique we use is a specific form of reversed subtitling: in reversed subtitling, the audio is presented in the viewers’ L1 while the subtitles are provided in the target language (TL). Although less common than standard subtitling (where audio is in the TL and subtitles in the L1), it has been argued to facilitate the acquisition of unfamiliar vocabulary (Danan, 1992; d’Ydewalle & Pavakanun, 1997). This method enables students to engage with content in their L1 while making connections between what they hear and the on-screen text (Talaván, 2012). In international advertisements, English is often employed in the audio. Given its status as a Lingua Franca, it is increasingly common to encounter it alongside less widely spoken languages, as in the present study involving Polish. Therefore, instead of referring to it as ‘reversed interlingual subtitling’ (L1 audio with TL subtitles), it is more accurate to describe it as ‘reversed multilingual subtitling’ (audio in a known language with subtitles in the TL).

4.2.2 Polish vocabulary test

The Polish vocabulary test was administered online and consisted of 15 target words (TW) and expressions appearing in the advert. We mainly focused on content words (nouns), instead of grammatical words, and very short phrases (e.g., *ono to jest*– it is–) as well as idiomatic expressions (e.g., *kocham cię*–I love you–). We avoided including cognates because they have already been shown to have a facilitating effect when learning from minimal input (e.g., Shoemaker & Rast, 2013) and their meanings could have been guessed just from the test. Therefore, we did not include transparent words in Spanish, Catalan, or English as TWs (e.g., *passport*–passport– or *pizama*–pyjamas–). All the TWs appeared once except for *kocham cię* and *ono to jest* (appearing twice), *cześć*–hi– (5 times) and *ja jestem*–I am– (6 times).

This multiple-choice meaning recognition test presented one word at a time to participants, who were then prompted to provide an answer before proceeding to the next item. Once an answer was submitted for a word, participants were unable to revisit previous items. Each TW (in Polish) was accompanied by three possible meanings (in English), along with the option “I don’t know” (please refer to Appendix A to see the test). The two distractors were items that had also appeared in the advert. In cases where the target was a multi-word item, there were also options consisting of multiple words (for instance, refer to items 6, 7, 9, 13 and 15 in Appendix A). To mitigate any recency effects, the order of appearance of the TWs was randomized. The entire test had a time limit of three minutes, with a maximum score of 15 points – one for each correct answer. The test assessed Level 3 knowledge, as per the intake typology proposed by Rast (2008). Specifically, it evaluated items that were perceived and comprehended, but not (re)produced.

4.2.3 Receptive vocabulary size test

V_YesNo v1.0 (Meara & Miralpeix, 2015) was used to measure participants’ receptive vocabulary size. It is an online test which aims to assess the number of words a person understands in English. It accomplishes this by presenting a series of words individually in a context-free environment. Test-takers are required to indicate whether they know the meaning of each word or not. The test covers words from the first ten frequency bands in English (1k to 10k) and provides an estimate of total receptive vocabulary size upon completion. To mitigate guessing, the test includes pseudo-words, ensuring that false alarms are factored into the final vocabulary size estimation. V_YesNo is derived from Yes/No tests initially proposed by Meara and Buxton (1987) and further developed by Meara and Jones (1988) for ~~English as a second language~~ learners. For a comprehensive understanding of the test’s functionality, the reader is referred to Meara and Miralpeix (2017, pp.113–133).

4.2.4 Questionnaire

The questionnaire comprised nine questions (refer to Appendix B), designed to collect participants' demographic details, their vocabulary learning strategies while engaging with the advert and their perceptions of this viewing experience. The initial question inquired participants about their proficiency in languages other than English, Spanish, and Catalan (the latter two being their native languages). Furthermore, to ensure participants were not familiar with Polish, the second question explicitly inquired about their familiarity with the TL and their proficiency level: beginner, intermediate, or advanced. Respondents who indicated familiarity with Polish were automatically excluded from the sample.

For the present study, several further questions are relevant: two open-ended questions prompted participants to reflect on their experience watching the Polish-subtitled advertisement and the vocabulary learning techniques they employed. The latter question facilitated the categorization of learning strategies in this study. Additionally, two further questions delved into participants' television viewing routines, specifically, whether they watched TV in Spanish, Catalan, or English. For those who indicated English, the questionnaire asked about their use of on-screen text and their subtitling preferences.

4.3 Procedure

Data from participants were collected in class by the teacher/researcher after obtaining the necessary permissions. Initially, participants in the EG were instructed to watch a short advertisement with the audio in English and subtitles in an unfamiliar language. They were directed to attempt to learn as much of this new language as possible. Following the first viewing, they were informed that they would watch the advertisement again for the same purpose, resulting in two viewings in total, that is, six minutes of exposure to Polish subtitles. Subsequently, participants were asked by the teacher to complete a vocabulary test available on the Moodle Virtual Campus of the course. They were informed that they had three minutes to answer all the questions and advised to select the "I don't know" option if uncertain, rather than guessing. Scores for this test were automatically recorded on the Virtual Campus of the course. Next, participants were provided with a questionnaire to complete, with no time limit specified, although they typically took an average of five minutes to answer it. Finally, they were debriefed about the results and explained how subtitled audiovisual input may contribute to learning new languages. The control group took the vocabulary test under the same conditions as the experimental group, but without watching the advert.

4.4 Data analysis

Test scores were processed and analysed using SPSS v.27. Similarly, responses to close-ended questions underwent systematic coding and analysis employing the same statistical software. Open-ended questions from the questionnaire were categorized and grouped *ad hoc* by the researchers and percentages calculated, to provide a description of participants' performance in RQs 2b and 3.

A preliminary independent samples *t*-test was first conducted to confirm there were no statistically significant differences between EG and CG in terms of English vocabulary knowledge (and that both would be equally able to understand the options provided in the Polish vocabulary test, see Table 1).

To answer RQ1, we analyzed how many questions were answered correctly in the Polish vocabulary test by participants in the EG after viewing the advert. We then compared these results with those of the CG, who had not seen the advert, using a *t*-test. This analysis helped ensure that any observed learning in the EG was attributable to the video exposure and not to random chance.

To address Research Questions 2a and 3, Chi-square tests were conducted. The continuous variable 'vocabulary scores' was categorized into three groups (low, middle, and high) using a *k*-means cluster analysis. In the case of RQ2 enquiring about a possible relationship between learning from the advert and variables such as the number of known languages or previous exposure to audiovisual input in other languages, Chi-square tests were conducted between vocabulary scores and these variables. To uncover the strategies employed by learners to facilitate vocabulary acquisition (RQ 2b), frequency counts were conducted to determine the prevalence of these strategies among novice learners.

Finally, in relation to RQ3, on how participants perceived the experience, several categories were derived from the data and a difference was made between 'positive' and 'negative' perceptions. A Chi-square test was performed to assess whether learners' perceptions had an impact on vocabulary scores.

5. Results

5.1 Vocabulary learning

Results on vocabulary learning from viewing showed that participants watching the subtitled advert were able to make form-meaning connections, recognizing the meaning of about half of the target Polish vocabulary after only six minutes of video exposure (see Table 2). These scores were also significantly higher than

those in the control group $t(133)=10.682, p<.001$, with a very large effect size ($d=2.24$; Plonsky & Oswald, 2014).

Table 2. Vocabulary scores in the EG and CG

	<i>M</i>	<i>SD</i>	Min.	Max.	95% CI
Experimental <i>n</i> = 106	7.33	2.43	2	14	[6.86-7.80]
Control <i>n</i> = 29	1.83	2.55	0	8	[0.86-2.80]

5.2 Variables influencing vocabulary learning

Results from the k-means cluster analysis indicated three distinct groups based on vocabulary scores: low ($n=30$, mean = 4.57, $SD=.152$), mid ($n=56$, mean = 7.60, $SD=1.15$) and high ($n=20$, mean = 10.70, $SD=1.08$). The Kruskal-Wallis test revealed a significant difference among the three groups ($\chi^2(2)=77.095, p<.001$) and the post-hoc analyses indicated that all groups differed significantly from each other.

Regarding proficiency in other languages, the ability to speak FLs other than English did not show a significant association with Polish scores ($X^2(2, 106)=.181, p=.913$). Similarly, familiarity with other Germanic languages such as German or Dutch, which were the most common languages learned after English, did not demonstrate a significant impact on Polish learning outcomes in either German ($X^2(2, 106)=1.312, p=.519$) or Dutch ($X^2(2,106)=1.610, p=.447$).

When assessing the possible influence of previous exposure to multimodal input and viewing habits on vocabulary scores, it was found that the frequency of exposure to original version TV (whether never, monthly, or weekly) was not related to Polish vocabulary scores ($X^2(4, 106)=.547, p=.969$). Similarly, the choice of subtitles in either L1 or the L2 (or the absence of subtitles) showed no significant effect. That is, no effects were observed in those setting the subtitles in Spanish ($X^2(2, 57)=.155, p=.926$), Catalan ($X^2(2, 57)=2.395, p=.302$), English ($X^2(2, 57)=.724, p=.696$); or in those that did not use subtitles at all ($X^2(2, 57)=.817, p=.665$).

5.3 Vocabulary learning strategies

Table 3 displays the strategies used by participants to learn vocabulary from the video, with cognitive strategies being the most prevalent. Notably, participants frequently mentioned reading subtitles in the new language and relying on visual

aids or their visual memory to foster retention. They also emphasized focusing on frequently repeated words (18.68%) and, to a lesser extent, on basic vocabulary (i.e., words that students tend to learn first when they are introduced to a new language, such as ‘thank you’, ‘I’, ‘dog’, etc.) (6.59%). Moreover, they also attempted to draw comparisons between the new vocabulary and words they knew in other languages. Additionally, the instruction to maximize learning and the opportunity to view the video twice proved beneficial in facilitating the acquisition of new vocabulary from the input.

Table 3. Vocabulary learning strategies and percentage of use

Strategies	Percentage	Strategy type*
Reading subtitles in Polish	26.37	C
Use of imagery/visual memory	23.08	C
Focus on repeated words	18.68	C
Cross-language comparisons	10.99	C
Intentional approach	8.79	M
Focus on basic words	6.59	C
Repeated viewing	3.30	M
Others	2.20	–


* C=Cognitive strategies, M=Metacognitive strategies

5.4 The viewing/learning experience

Based on participants’ responses (please refer to Table 4), 59.82% of the sample perceived the experience negatively, describing it as very difficult (resulting in feelings of getting lost), weird, challenging, or stressful. However, 40.18% of the participants perceived the task as fun, interesting, and useful for learning new words, finding it easy to navigate. Results from the Chi-square test revealed that perceiving the experience as being ‘positive’ or ‘negative’ did not impact vocabulary scores ($X^2(2, 106) = .387, p = .824$).

6. Discussion

Findings from the study make evident that making form-meaning connections for new vocabulary in a new language is possible with just very minimal information (i.e., a subtitled advert watched twice, six minutes of exposure in total). It should be considered that the audio was in a language participants knew, even if it was

 Table 1. Participants' perceptions of the viewing/learning experience

The viewing/learning experience	Percentage	Experience perception *
Very difficult, I got lost	26.79	–
Strange, weird	21.43	–
Fun, interesting	18.75	+
Useful to learn new words	16.07	+
Challenging	8.04	–
Easy	5.36	+
Stressful	3.56	–

* +: Positive experience, –: Negative experience.

not their L1: this is a common scenario learners may encounter due to the widespread use of English as a *lingua franca* in mass media. The fact that the English vocabulary present in the advert was within the participants' level in this language made input easy to follow, minimizing the effects of not having been exposed to L1 audio. Additionally, participants were instructed to deliberately try to learn the new language, which did not happen in Bisson et al. (2014), possibly causing the difference in the two studies' outcomes. Therefore, it is probable that task orientation changed participants' focus, as suggested by VanPatten (2014), and prompted learners to make an extra effort to commit words to memory (Schmidt, 2001). As the video was shorter (it was an advert instead of an episode in Bisson et al. 2014), participants did not need to maintain attention for an extended duration, facilitating retention. Last, the fact that participants watched the same advert twice might have also assisted the learning of some TWs since learners could easily (dis)confirm the hypotheses already formed during the first viewing, in line with previous studies which have shown that repeated viewing is beneficial for language development (Muñoz et al., 2022).

Furthermore, it is noteworthy that despite similar percentages of participants perceiving the task as 'difficult/strange' and 'interesting/useful', vocabulary acquisition did not appear to be contingent on learners' perceptions. This suggests that first of all, learning occurs regardless of how learners perceive the task (similarly to Pattemore et al., 2024, where perceptions of learning in *en route* learners did not necessarily align with the actual learning). Also Sydorenko (2010) found significant vocabulary gains from short video viewing, even if many participants had not perceived any learning from the treatment. Secondly, other factors, such as individual differences, may play a more significant role in explaining learning outcomes. For instance, Miralpeix et al. (2023) observed aptitude effects when learning from visual stimuli, and Rast et al. (2014) indicated that factors such as

working memory, learning style, and motivation were associated with task performance in Polish during the initial five hours of language contact in a classroom setting.

Regarding variables that may affect linguistic gains from the subtitled video, previous knowledge of other languages did not affect learning at FE. This may appear contradictory to findings by Gibson and Hufeisen (2003) or Hirosh and Degani (2018) in *en route* learners, or by Rast (2010a) in novice learners, suggesting an advantage for multilinguals when learning new languages. However, it is important to note that we did not include cognates in our test, and our participants had no knowledge of any Slavic language which was closer to Polish (they were bilingual in two Romance languages and were studying English, which belongs to the Germanic branch). It should also be noted that if being bilingual could confer a lexical advantage in the early stages of language learning (e.g., Hopp et al., 2019), all participants in our study were bilingual. This means that they all would have an advantage in an FE experience compared to monolinguals.

Despite the non-significant results of the Chi-square test, participants indicated in the questionnaire that they were using other known languages for cross-linguistic comparisons, although this strategy might not have been helpful in achieving higher scores on the vocabulary test, as no cognate words were included. In connection with this, while the four primary languages examined in the study (Catalan, Spanish, English, and Polish) all employ the Roman alphabet, Polish exhibits a slightly different orthography compared to the other three. This includes the presence of diacritics unique to Polish, which consequently increases the learning challenge for the TWs and diminishes the effectiveness of cross-language comparisons. For example, in Polish we find the *kreska* (graphically similar to the acute accent) on the letters ć, ń, ó, ś, ź; the *kropka* (overdot) in the letter ż; the stroke in the letter ł; or the *ogonek* (little tai) in the letters ą, ę. Similarly, Polish features a syllable structure characterized by numerous consonant clusters not commonly found in the other three languages, e.g., cz, dz, sz (Orzechowska & Wiese, 2015).

Regularly watching TV in the original version, with or without subtitles, did not exhibit any correlation with learning outcomes in this study. At first glance, this may appear contradictory to Vanderplank's (2016) findings among *en route* learners, where individuals with more viewing experience tended to adopt more strategic approaches, such as making connections, comparisons, or corrections more frequently. However, exposure to audiovisual input in a FL may not automatically lead to better performance when watching videos in a completely new language, at least not within the initial moments of contact, as each new language poses its own challenges. It is possible that more time and engagement are needed for learners to get accustomed to this practice in the new language.

It is interesting to observe that the vocabulary learning strategies reported by participants in the questionnaire can be correlated with the steps students undertake when acquiring new words, as proposed by Brown and Payne (1994), also referenced in Hatch and Brown (1995). Consequently, we can categorize the actions students take when intentionally learning through viewing at FE according to the stages outlined by Brown and Payne (1994), as depicted in Figure 1. Among the five stages they propose, we will concentrate on three: ‘forming a clear mental representation of the new word forms’, ‘grasping the meaning of words’, and ‘establishing a memory connection between form and meaning’. We will exclude ‘accessing sources for discovering new words’, which would precede the aforementioned three stages, as participants were already provided with the video source to extract lexical information. Similarly, we will omit the fifth and final stage, ‘using the words’, as our study did not evaluate production.

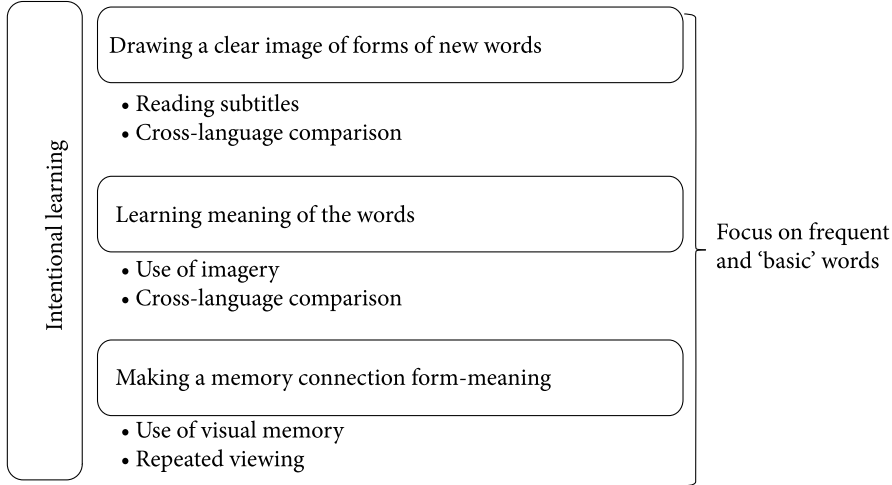


Figure 1. Vocabulary learning strategies at FE related to the steps students take when learning new words (adapted from Brown & Payne, 1994)

Therefore, participants, to form a clear mental representation of the new word forms in this type of reversed subtitling condition, would read the subtitles (a discovery strategy, according to Schmitt, 1997) and possibly engage in cross-language comparisons. Progressing to the next stage (i.e., learning the meaning of new words), they would employ imagery and potentially rely on other languages they are proficient in, seeking similarities (as in Sydorenko, 2010). In the final stage, when establishing a strong connection between the form and the meaning of words, they would use their visual memory, and repeated viewing would aid in solidifying these connections (Muñoz et al., 2022). Furthermore, at all stages, there

would be a focus on frequently encountered items and fundamental words (as confirmed by the participants' responses to the questionnaire). This approximation helps us to understand how learners confront the challenge of 'breaking into the wild' when exposed to multimodal input in an unfamiliar language, helping us to see how they possibly make sense of such input and try to turn it into intake.

7. Conclusion, limitations, and further research

The present study confirms that intentional vocabulary learning occurs even with minimal exposure to a completely new language through subtitled videos, particularly when the audio language is familiar, facilitating the establishment of a form-meaning connection. Our findings align with Gullberg et al. (2010, 2012), underscoring adults' capacity to learn under minimal input conditions. Thus, short, subtitled videos can serve as a tool to promote lexical acquisition in novice learners, aiding in the development of vocabulary from the outset. Optimal effectiveness may be achieved when the content is engaging and includes repeated vocabulary. Furthermore, starting with audio in a language learners know (not necessarily the L1, as in this case) can promote fast-mapping, while incorporating audio in the new language may facilitate recognition of written forms, as evidenced in studies with *en route* beginner learners (see, for example, Casulleras and Miralpeix, Chapter 1, this volume).


The present study has demonstrated that learning occurs regardless of task perception (whether perceived as 'easy' or 'difficult'). While direct effects of knowledge of other languages on learning were not evident, cross-language comparisons were acknowledged to take place by participants, who were already multilingual. Additionally, we have observed variations in the strategies employed to learn new vocabulary compared to those used in other contexts or with different input types. This insight holds significant value for educators aiming to support learners during their initial language acquisition stages, especially when devising instructional sequences that incorporate video viewing. Offering guidance on how to use these learning strategies can enrich the prospects for novel learning opportunities.

Addressing the limitations, it is important to note that only one specific combination of languages was investigated in this study, with a limited number of TWs and immediate testing (assessing only one intake type). To draw more robust conclusions regarding learning from viewing at FE, studies incorporating different language combinations, a broader spectrum of vocabulary, and various intake types (such as word form recognition and recall, word meaning recall, etc.) should be conducted. In addition, the intentional learning approach implemented

in the current design undoubtedly influenced the processing of TWs. Therefore, it would be valuable to investigate whether novice learners also acquire new vocabulary incidentally, through exposure to multimodal input during FE.

Furthermore, data-driven classifications, not predetermined by researchers, were used to identify patterns in the data. However, alternative approaches, such as providing students with a pre-piloted list of potential strategies and prompting them to indicate which strategies they employed (e.g., using plot clues, identifying word categories, linking words to scenes or characters, recalling entire subtitles to make inferences, etc.), could have been employed. This approach would have enabled learners to recall actions they engaged in while watching the video, but which might not have been mentioned otherwise. Additionally, a different approach, as suggested by Meara (2009), could involve uncovering groups of participants who adopt similar strategies rather than focusing solely on strategies that frequently co-occur. We propose conducting this 'by-subjects' analysis in future research on the topic. Finally, it is recommended to improve our understanding of the roles played by individual differences (e.g., aptitude or working memory) and to advance our exploration of learning at FE through eye-tracking techniques. This will enable us to investigate potential correlations between processing and learning in these contexts.

Funding


This work was supported by grant PID2019-110594GB-I00 from the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/501100011033) 














Acknowledgements














We thank Neus Frigolé for her help with data collection, Paulina Olender for her revision of the Polish test, as well as the students who participated in the study. We are also very grateful to the chapter reviewers for their feedback and suggestions, which helped improve our work.














References

Allegro (2016, Dec.). *English for beginners*. Allegro. Retrieved on 28 May 2024 from <https://rb.gy/9phftk>

 Bisson, M. J., Van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, 35(2), 399–418.

-  Borrás, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *Modern Language Journal*, 78(1), 61–75.
- Brown, C., & Payne, M. E. (1994, March 8–12). Five essential steps of processes in vocabulary learning [Conference presentation]. 28th TESOL Annual Convention, Baltimore, MD. 
- Carroll, S. E. (2014). Processing ‘words’ in early-stage foreign language acquisition: A comparison of first exposure and low proficiency learners. In Z. H. Han & R. Rast (Eds.), *First exposure to a second language* (pp. 107–138). Cambridge University Press.
- Cobb, T. (n.d.). VocabProfile VP Classic [Computer software]. Retrieved on 28 May 2024 from <https://www.lexutor.ca/vp/eng/>
-  Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. *Brain Research*, 1123(1), 168–178.
-  Danan, M. (1992). Reversed subtitling and Dual Coding Theory: New directions for foreign language instruction. *Language Learning*, 42(4), 497–527.
-  De Diego Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PloS ONE*, 2(11).
- De Groot, A. M. B., & van Hell, J. G. (2005). The learning of foreign language vocabulary. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 9–29). Oxford University Press.
-  De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020). Learning English through out-of-school exposure: Which levels of language proficiency are attained and which types of input are important? *Bilingualism: Language and Cognition*, 23(1), 171–185.
-  Dimroth, C., Rast, R., Starren, M., & Watorek, M. (2013). Methods for studying the acquisition of a new language under controlled input conditions. *Eurosla Yearbook*, 13, 109–138.
-  d’Ydewalle, G., & Pavakanun, U. (1997). Could enjoying a movie lead to language acquisition? In P. Winterhoff-Spurk & T. H. A. van der Voort (Eds.), *New horizons in media psychology: Research cooperation and projects in Europe* (pp. 145–155). Springer.
-  Folia, V., Uddén, J., de Vries, M., Forkstam, C., & Petersson, K. M. (2010). Artificial language learning in children. *Language Learning*, 60(Suppl. 2), 188–220.
- Gibson, M., & Hufeisen, B. (2003). Investigating the role of prior foreign language knowledge: Translating from an unknown into a known foreign language. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *The multilingual lexicon*. Springer. 
-  Griffiths, C. (2008). Strategies and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 83–98). Cambridge University Press.
-  Gu, P. Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46(4), 643–679.
- Gullberg, M., & Indefrey, P. (Eds.) (2010). The earliest stages of language learning. *Language Learning*, 60(Suppl. 2).
-  Gullberg, M., Roberts, L., & Dimroth, C. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *IRAL – International Review of Applied Linguistics in Language Teaching*, 50(4), 239–276.

-  Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., & Indefrey, P. (2010). Adult language learning after minimal exposure to an unknown natural language. *Language Learning*, 60(Suppl. 2), 5–24.
-  Han, Z. H., & Rast, R. (Eds.), (2014). *First exposure to a second language*. Cambridge University Press.
- Hatch, E., & Brown, C. (1995). *Vocabulary, semantics, and language education*. Cambridge University Press.
-  Hirosh, Z., & Degani, T. (2018). Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin and Review*, 25(3), 892–916.
-  Hofweber, J. E., Aumonier, L., Janke, V., Gullberg, M., & Marshall, C. (2022). Breaking into language in a new modality: The role of input and individual differences in recognising signs. *Frontiers in Psychology*, 13.
-  Hofweber, J. E., Aumonier, L., Janke, V., Gullberg, M., & Marshall, C. (2023). Which aspects of visual motivation aid the implicit learning of signs at first exposure? *Language Learning*, 73(S1), 33–63.
-  Hopp, H., Vogelbacher, M., Kieseier, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction*, 61, 99–110.
-  Kellerman, E. (1979). Transfer and non-transfer: Where we are now. *Studies in Second Language Acquisition*, 2(1), 37–57.
-  Kusyk, M., & Sockett, G. (2012). From informal resource usage to incidental language acquisition: Language uptake from online television viewing in English. *ASp, la revue du Geras* 62, 45–65.
-  McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, 7(7), 702–704.
- Meara, P.M. (2005). *LLAMA language aptitude tests*. Lognostics. Retrieved on 28 May 2024 from www.llamatests.org
-  Meara, P.M. (2009). Review of ‘Vocabulary learning strategies and foreign language acquisition’. *System*, 37(3), 545–547.
-  Meara, P.M., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing*, 4(2), 142–154.
- Meara, P.M., & Jones, G. (1988). Vocabulary size as a placement indicator. In P. Grunwell (Ed.), *Applied Linguistics in Society* (pp. 80–87). CILT.
- Meara, P.M., & Miralpeix, I. (2015). V_YesNo. Lognostics. Retrieved on 28 May 2024 from https://www.lognostics.co.uk/tools/V_YesNo/V_YesNo.htm
- Meara, P.M., & Miralpeix, I. (2017). Tools for researching vocabulary. *Multilingual Matters*.
-  Miralpeix, I., Gesa, F., & Suárez, M. M. (2023). Vocabulary learning from subtitled input after minimal exposure. In B. Reynolds (Ed.), *Vocabulary learning in the wild* (pp. 263–283). Springer.
-  Muñoz, C., Pattemore, A., & Avello, D. (2022). Exploring repeated captioning viewing as a way to promote vocabulary learning: Time lag between repetitions and learner factors. *Computer Assisted Language Learning*.

-  Orzechowska, P., & Wiese, R. (2015). Preferences and variation in word-initial phonotactics: A multi-dimensional evaluation of German and Polish. *Folia Linguistica*, 49(2), 439–486.
-  Pattemore, A., Suárez, M.-M., & Muñoz, C. (2024). Perceptions of learning from audiovisual input and changes in L2 viewing preferences: The roles of on-screen text and proficiency. *ReCALL*, 36(2), 135–151.
-  Pavičić Takač, V. (2008). *Vocabulary learning strategies and foreign language acquisition*. Multilingual Matters.
-  Plonsky, L., & Oswald, F. L. (2014). How big is ‘big’? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912.
-  Rast, R. (2008). *Foreign language input: Initial processing*. Multilingual Matters.
-  Rast, R. (2010a). The role of linguistic input in the first hours of adult language learning. *Language Learning*, 60(Suppl. 2), 64–84.
-  Rast, R. (2010b). The use of prior linguistic knowledge in the early stages of L3 acquisition. *IRAL – International Review of Applied Linguistics in Language Teaching*, 48(2–3), 159–183.
- Rast, R., Watorek, M., Hilton, H., & Shoemaker, E. (2014). Initial processing and use of inflectional markers: Evidence from French adult learners of Polish. In Z. H. Han & R. Rast (Eds.), *First exposure to a second language* (pp. 64–106). Cambridge University Press.
-  Reynolds, B. L. (Ed.), (2023). *Vocabulary learning in the wild*. Springer.
-  Reynolds, B. L., Cui, Y., Kao, C. W., & Thomas, N. (2022). Vocabulary acquisition through viewing captioned and subtitled video: A scoping review and meta-analysis. *Systems*, 10(5), 133.
-  Ristin-Kauffman, N., & Gullberg, M. (2014). The effects of first exposure to an unknown language at different ages. *Bulletin Suisse de Linguistique Appliquée*, 99, 17–29.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199–227). Cambridge University Press.
- Schmitt, N. (2000). Issues in the emerging area of vocabulary learning strategies. In J. Arabski (Ed.), *Studies in foreign language acquisition and teaching* (pp. 107–116). Wydawnictwo Uniwersytetu Śląskiego.
-  Shoemaker, E., & Rast, R. (2013). Extracting words from the speech stream at first exposure. *Second Language Research*, 29(2), 165–183.
- Singleton, D., & Little, D. (1984). A first encounter with Dutch: Perceived language distance and language transfer as factors in comprehension. In L. Mathuna & D. Singleton (Eds.), *Language across cultures* (pp. 259–270). Irish Association for Applied Linguistics.
-  Stewart, M. A., & Pertusa, I. (2004). Gains to language learners from viewing target language closed-captioned films. *Foreign Language Annals*, 37(3), 438–442.
- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, 14(2), 50–73. 10125/4421 

Talaván, N. (2012). Justificación teórico-práctica del uso de los subtítulos en la enseñanza-aprendizaje de lenguas extranjeras. *TRANS. Revista de Traductología*, 16, 23–37. Retrieved on 28 May 2024 from: http://www.trans.uma.es/trans_16/Trans16_023-037.pdf

Thompson, I., & Rubin, J. (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals*, 29(3), 331–342.

Vanderplank, R. (2016). Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing and subtitles for language learning. *Palgrave Macmillan*.

VanPatten, B. (2014). Epilogue: Input processing by novices – Issues in the nature of processing and in research methods. In Z. H. Han & R. Rast (Eds.), *First exposure to a second language* (pp. 193–207). Cambridge University Press.

Zhang, B., & Li, C. (2011). Classification of L2 vocabulary learning strategies: Evidence from exploratory and confirmatory factor analysis. *RELC Journal*, 42(2), 141–154.

Appendix A. Meaning recognition vocabulary test

Target Word (TW) *	Option A	Option B	Option C	Option D
1. chleb (1)	breakfast	bread	fridge	I don't know
2. dziękuję ci (1)	goodbye	you are perfect	thank you	I don't know
3. dziadkiem (1)	wall	beach	grandpa	I don't know
4. ręcznik (1)	fork	bath	towel	I don't know
5. kapcie (1)	towel	pyjamas	slippers	I don't know
6. ono to jest (2)	it is	you are perfect	I love you	I don't know
7. szczoteczka do zębów (1)	toothbrush	I love you	I'm gonna fucking kill you	I don't know
8. śniadanie (1)	knife	breakfast	bread	I don't know
9. kocham cię (2)	you are perfect	I am gonna fucking kill you	I love you	I don't know
10. początkujących (1)	pyjamas	beginners	fork	I don't know
11. psem (1)	dog	beach	grandpa	I don't know
12. drogę (1)	toilet	mirror	way	I don't know
13. bądź dobrym (1)	can you show me	thank you	be good	I don't know
14. cześć (5)	hi	thank you	be good	I don't know
15. ja jestem (6)	I am	you are	you can	I don't know

* Correct answers for each TW appear in bold and TW frequency in the advert within parenthesis.

Appendix B. Questionnaire

1. Apart from Spanish, Catalan and English, do you know any other language(s)?
☐ Yes ☐ No
 If you said 'yes', which level do you have in each? (Beginner, Intermediate, Advanced)
 Language: Level:
 Language: Level:
 2. Before watching the clip, did you know any Polish? ☐ Yes ☐ No
 If you said 'yes', which level do you have in Polish?
☐ Beginner ☐ Intermediate ☐ Advanced
 3. How was the experience of watching an advert in English subtitled in Polish? Please describe it.
 4. Did you notice that the language in the subtitles was Polish? ☐ Yes ☐ No
 5. 'Did you expect to learn anything from the new language when seeing the advert'? ☐ Yes ☐ No
 If you said 'yes', what?
☐ grammar ☐ vocabulary ☐ morphology ☐ pronunciation ☐ fixed expressions
☐ others (please specify):
 6. You have the result of the quiz. How do you feel about it?
 - I have learned what I expected ☐
 - I have learned more than I expected ☐
 - I have learned less than I expected ☐
 7. Which strategies did you use to help you to learn new words?
 8. Which language do you usually watch TV series, films and/or videos on YouTube in:
☐ Spanish ☐ Catalan ☐ English
☐ Others? _____
 9. Do you usually watch TV series, films and/or videos in English? ☐ Yes ☐ No
 If you said 'yes':
 How often? ☐ Less than once a month
 ☐ Between 1–3 times / month
 ☐ Between 1–3 times / week
 ☐ Between 4–6 times / week
 ☐ Every day
- Please select the language you set the subtitles in:
☐ Spanish ☐ Catalan ☐ English ☐ No subtitles at all



Author Queries

- Please provide a complete reference for the citation '(Allegro, 2008)' in this article.
- Please provide a citation for the reference id "c9-CIT0001 (Allegro (2016))" since citation is missing in the article.
- Please provide the correct DOI for the reference "c9-CIT0055 (Sydorenko, T. (2010))".

(Allegro, 2008) in the text should be (Allegro 2016), as in the references.

Correct DOI for Sydorenko 2010 has been provided.

Thank you!