

The Influence of Emotional and Foreign Language Context in Content Learning

Candice Frances^{1*}

Angela De Bruin^{1,2}

Jon Andoni Duñabeitia^{3,4}

¹ *BCBL, Basque Center on Brain, Language and Cognition; Donostia, Spain*

² *Department of Psychology, University of York, York, United Kingdom*

³ *Centro de Ciencia Cognitiva – C3, Universidad Nebrija*

⁴ *Department of Language and Culture, The Arctic University of Norway*

*Contact information:

Candice Frances

c.frances@bcbl.eu

Basque Center on Cognition, Brain and Language (BCBL)

Paseo Mikeletegi 69, 2nd floor,

20009 Donostia - Spain

+34 943309300

Abstract

Prior research has found reduced emotionality with foreign language use, especially with single words, but what happens if emotionality is conveyed throughout a longer text? Does emotionality affect how well we remember and associate information, i.e., content learning? We played participants descriptions of two invented countries and tested how well they remembered facts about these countries. Each participant listened to one positive and one neutral description, which was read either in their native language (Spanish) or in their foreign language (English). Participants remembered facts they heard in positive semantic contexts better than those learned in neutral semantic contexts, and did better in their native than their foreign language. Importantly, there was no interaction between language and emotionality, suggesting that the previously reported decrease in emotionality in a foreign language might not extend to all areas of foreign language use. **Words:** 139

Keywords: emotionality; foreign language effects; non-native languages; learning; auditory modality

The Influence of Emotional and Foreign Language Context in Content Learning

As study abroad programs become more common, it is imperative that we understand how foreign languages (FL) affect our learning. For example, are we able to learn new content in a FL to the same extent as in our native language (NL)? There is a substantial amount of literature assessing this question in children, but there is little published research regarding adult learning. Furthermore, the current adult literature focuses mostly on memory for single words (e.g., Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Caldwell-Harris, 2009; Ferre, Garcia, Fraga, Sanchez-Casas, & Molero, 2010). One possible mechanism for improving content learning in an FL—drawing from the NL literature—is using emotionality to enhance memory. Emotional items are easier to remember in our NL than in our FL (see Caldwell-Harris, 2014 for a review). But, can this strategy be used to improve performance in an FL? Importantly, prior single word research has found reduced emotionality effects in an FL, but what happens if emotionality is conveyed throughout a longer text rather than in single words? The current study attempts to expand on these questions, testing memory for information embedded in an emotional context, to see whether this can boost content learning in an FL.

One of the most common types of programs that use FL to teach new information is content and language integrated learning (CLIL). CLIL refers to a curriculum-based approach used to teach content courses using a second language, in order to teach both content and language through immersion. Although research on the language learning aspects of CLIL quite conclusively shows an improvement in FL use and comprehension (Admiraal, Westhoff, & De Bot, 2006; Aguilar & Rodríguez, 2012; Bergroth, 2006; Dalton-Puffer, 2007; Jiménez Catalán & Ruiz de Zarobe, 2009; Ouazizi, 2016; Serra, 2007; Xanthou, 2011; although see Dallinger, Jonkmann, Hollm, & Fiege, 2016 for no improvement), the research on content learning is less clear-cut (Dalton-Puffer, 2011). There are studies that find positive effects (Day & Shapson, 1996; Jäppinen, 2005; Ouazizi, 2016; Pérez Cañado, 2018; Surmont, Struys, Van Den Noort, & Van De Craen, 2016; Van de Craen, Ceuleers, & Mondt, 2007; Xanthou, 2011), while others find negative (Anghel, Cabrales, & Carro, 2016; Dallinger et al., 2016; Fernández-Sanjurjo, Fernández-Costales, & Arias Blanco, 2017) or null effects (Admiraal et al., 2006; Bergroth, 2006; Serra, 2007; Stohler, 2006). Consequently, these results paint a less than clear picture of how children learn new content in an FL.

The literature on adult FL-medium learning is more limited, with most of the reported benefits being associated with language (e.g., Yang, 2014) and not content. These studies often show no difference between the control and experimental group in overall performance at the end of the course (e.g., Hernandez-Nanclares & Jimenez-Munoz, 2015), but very few examine the immediate understanding and learning of new content in an FL. Those that do report a difference find that instruction in an FL is detrimental, particularly without FL support (Roussel, Joulia, Tricot, & Sweller, 2017). These results have been accounted for in the context of cognitive load theory, which suggests a working memory overload for individuals trying to learn content in a language they are not proficient in (Roussel et al., 2017). Importantly, contributing to this literature would influence and possibly improve teaching methods for adults studying in an FL.

Given the difficulties in learning new content in an FL, we need to find ways of compensating for or aiding in improving performance. One way of doing this is by applying what we know from the NL studies. Considering this literature, one of the variables that aids learning is emotionality, as learning emotional words (see Caldwell-Harris, 2014 for a review), or seeing neutral words in emotional contexts (Erk et al., 2003; Erk, Martin, & Walter, 2005), improves memory performance. However, several studies show that speakers are less emotional in an FL than in an NL context (Dewaele, 2010; Harris, Gleason, & Ayçiçeği, 2006; Pavlenko, 2002). One might extrapolate from these studies that

using emotionality as a tool to boost learning would not be as efficient in an FL. Indeed, Anooshian and Hertel (1994) found that participants remembered emotional words better than neutral words in their NL, but not in their FL. This is in line with foreign language effect (FLE) research supporting a reduction in emotionality in an FL (Costa, Foucart, Hayakawa, et al., 2014; Costa, Foucart, Arnon, Aparici, & Apesteguia, 2014; Costa, Vives, & Corey, 2017; Hadjichristidis, Geipel, & Savadori, 2015; Keysar, Hayakawa, & An, 2012, but see Vives, Aparici, & Costa, 2018). Conversely, other studies find the same effects of emotion on memory in both languages (Ayçiçeği & Harris, 2004; Caldwell-Harris, 2009; Ferré, Ventura, Comesaña, & Fraga, 2015; Ponari et al., 2015). Therefore, it is not clear how the effects of emotionality in an FL compare to those of the NL.

Nevertheless, these conflicting results may be explained by alternative accounts, such as a reduction in intuitive responses and depletion of cognitive resources (Geipel, Hadjichristidis, & Surian, 2015a, 2015b, 2016) or triggering of different cultural norms (Gawinkowska, Paradowski, & Bilewicz, 2013) in the FL. Gawinkowska et al. (2013) suggest that the FLE is due to a difference in social and cultural norms rather than a difference in emotional impact between languages. Regardless of the origin of the effect, it is not clear whether people respond similarly to emotional stimuli in their NL and FL, nor whether they benefit from the effects of emotionality on memory the same way in an FL as in an NL. Furthermore, the paradigms used thus far predominantly focus on emotionally-charged words in isolation rather than in context (e.g., Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Caldwell-Harris, 2009; Ferre et al., 2010) and are limited to using single-word auditory material. This is particularly relevant since, contrary to this approach, information taught in classrooms is most commonly conveyed in context.

The objective of this study is to investigate content learning and how it is affected both by an FL and an emotional context. There is little research directly comparing acquisition of new concepts and knowledge in a bilingual's NL and FL. Likewise, there is no research looking into the effects of emotionality in this context, nor listening to texts manipulating emotional context semantically. Understanding how these variables interact can contribute to classrooms that use an FL as the medium of teaching, improving methods and efficacy. To address this, we had participants listen to two descriptions of countries (one positive and one neutral) in either their NL (Spanish) or an FL (English), followed by a multiple choice test. Using longer texts than those used in prior research, we aimed to create a more realistic replication of information processing and acquisition. Thus, participants were required to learn interrelated facts that made a coherent whole, rather than independent pieces of information disconnected from each other (see Frances, de Bruin, & Duñabeitia, n.d., for a similar study using vocabulary learning and non-related information). This would allow them to create more complex networks of meaning, which in turn would allow us to understand how semantic context can affect memory for individual facts within these larger conceptual networks. We hypothesized that despite the fact that their overall performance was likely to be poorer in the FL than in the NL contexts, bilinguals would not show an FLE, but instead would present similar emotionality effects in both languages. The rationale for this is that, if the FL affects responding by reducing reliance on intuition or simply requires more cognitive resources—as suggested before—the effect of emotionality should remain the same.

Methods

Participants

Participants were 76 native Spanish speakers (38 in each language group, 9 male, $M_{\text{age}} = 33.86$, $SD_{\text{age}} = 9.14$), recruited through language schools and randomly assigned to either the NL or FL context. All participants completed a test of English vocabulary (LexTALE; Lemhöfer & Broersma, 2012) and had

a minimum score of 60%. This is equivalent to a minimum of a B2 level according to the Common European Framework of reference for languages, with 50 participants at the B2 level range and 26 at the C1/C2 level (Lemhöfer & Broersma, 2012). Participants in the two language contexts were matched on age and education level (i.e., highest level of schooling achieved, in all cases at least high school) according to the sociodemographic information gathered, as well as multiple language variables. They were asked to rate their English level overall on a 1-to-10 scale as well as their listening, reading, speaking, and writing skills in that language. They also reported their estimated age of acquisition of English and the amount of time spent living in an English speaking country ($M = 3.08$ months $SD = 4.65$ months; all were living in Spain at the time of testing). Finally, they were matched on English and Spanish vocabulary knowledge as assessed by LexTALE (Lemhöfer & Broersma, 2012) and the LexTALE-Esp (Izura, Cuetos, & Brysbaert, 2014). For a summary of these variables, see Table 1 and on-line supplementary materials for means, distributions, and Bayes factors. The study and protocols were approved by the ethics committee at the BCBL.

Instruments

We created the description for two imaginary countries including 50 different items of information (e.g., national sport and population—see on-line supplementary materials for the list of test items). These two descriptions were then modified with filler sentences to include a more positive or neutral description of the country (e.g., neutral: “The population of Tecamer is defined politically as left wing, although they are considered generally quite moderate in their political, economic, and social opinions” and positive: “The population of Tecamer is defined politically as left wing and supports freedom, tolerance, and social inclusion as well as equal opportunity, leading many campaigns against discrimination”). The Spanish and English versions were created simultaneously and were matched on length. The texts were 50 to 56 sentences long and the average number of words in the English and Spanish versions were matched (1278.5 and 1317, respectively). The two emotional conditions were matched within languages on lemmatized word frequency of the content words (Spanish using LEXESP database, Sebastián-Gallés, Martí, Carreiras, & Cuetos, 2000; English using the HAL database, Lund & Burgess, 1996—Table 2). Importantly, the positive and neutral versions of the texts significantly differed on the mean valence and arousal of the words used, according to the ANEW database (Bradley & Lang, 1999) (valence: $BF_{01} = 2.42e+11, 5.22e-18$; arousal: $BF_{01} = 3.068e+10, 4.14e-17$). The number of high arousal (arousal >5) and high valence (valence >5) words also varied by condition (6% of the neutral condition and 12% of the positive condition was high valence word—see Appendix).

These four texts (2 countries, each with a neutral and a positive version) were read aloud and recorded by four female native Spanish speakers and four female native English speakers. Each recording lasted between 6.85 and 8.07 minutes ($M_{\text{duration}} = 7.51$ minutes, $SD_{\text{duration}} = .333$ minutes).

Procedure

Participants accessed the experiment through LimeSurvey (Schmitz, 2019). First, they filled out a demographics and language questionnaire and then listened to two audio files, one of each country in a given emotionality and different speakers (out of the 4 possible ones in that language). Each participant heard recordings in only one language and carried out the rest of the study in that same language. The order of the countries, emotional condition, and emotional condition/country matching were all randomized across participants to avoid any strategic or order effects. Once participants finished listening to the audio files, they proceeded to answer 50 multiple-choice questions about the stimuli content. These questions had 4 answer choices and participants were asked to pick one for each of the countries.

Analysis

The size of the sample was determined using GPower (Faul, Erdfelder, Lang, & Buchner, 2007), assuming a small to medium size interaction ($\eta_p^2 = .05$) and 95% power.

We carried out a two-way mixed ANOVA exploring the effects of emotionality and language on performance in the test to address whether performance was better in the NL or FL, whether emotional semantic context affects performance, and whether there was an interaction between the two. A main effect of language would indicate whether participants perform better in one of their languages, whilst a main effect of emotionality would reveal whether the emotional manipulation affected performance. Finally, any interaction between language and emotionality would show whether the effect of emotionality is modulated by language—meaning, emotionality affects people differently in the FL than the NL. In all cases, assumptions of statistical tests were met.

We followed these tests up with Bayesian factors (Jeffreys, 1961), which represent the likelihood of one model—in this case, the null hypothesis—over another—in this case, the alternative hypothesis. For example, a BF_{01} of 5 means that the null hypothesis is 5 times more likely to be true than the alternative one and a BF_{01} of .2 means that the alternative hypothesis is 5 times more likely to be true than the null. These Bayes Factors have become increasingly common as an alternative to frequentist models (Poirier, 2006), in particular for ANOVAs (Rouder, Morey, Speckman, & Province, 2012).

Results

First, we calculated the internal consistency between the questions of each country and found that the tests had good internal consistency (Mufelo $\alpha = .84$; Tecamer $\alpha = .86$).

We removed participants who were outliers, meaning 1.5 IQR away from the median in either condition (positive or neutral) for each language group. Using this procedure, we removed one participant from the English group and 3 from the Spanish group. The same tests were carried out with and without the outliers and the results were consistent between the two.

We carried out a two-way mixed ANOVA with emotionality and language on performance on the test (see Table 3 for means, standard deviations, and 95% confidence intervals). There was a significant main effect of emotionality, such that participants performed better in the positive ($M = 69.00\%$, $SD = 13.95\%$) than the neutral condition ($M = 65.97\%$, $SD = 14.71\%$), $F(1,70) = 8.54$, $p = .005$, $\eta_p^2 = .109$, $BF_{01} = .146$, $\text{error}\% = 1.26 \times 10^{-6}$ (see Figure 1 and on-line supplementary materials). There was also a main effect of language, such that participants performed better in their NL (Spanish: $M = 74.6\%$, $SD = 11.2\%$) than in their FL (English: $M = 60.3\%$, $SD = 11.6\%$), $F(1,70) = 26.83$, $p < .001$, $\eta_p^2 = .277$, $BF_{01} = 1.40 \times 10^{-4}$, $\text{error}\% = 1.29 \times 10^{-7}$. There was no interaction between the two factors, $F(1,70) = .104$, $p = .748$, $\eta_p^2 = .001$. A Bayesian repeated measures ANOVA comparing the model with the interaction (emotionality * language) and without the interaction term confirmed that there was moderate evidence that the addition of the interaction term led to an equally likely model, $BF_{01} = 4.12$, $\text{error}\% = 3.15$ —namely, no interaction was over 4 times more likely than an interaction. We also ran an independent samples t-test on the emotionality effect—namely the score on the positive condition minus the score on the neutral one for each of the language conditions—and again found moderate evidence in support of the null hypothesis, $BF_{01} = 3.93$, $\text{error}\% = .012$.

Discussion

In the current study, we addressed the questions of whether learning new information in an FL could be improved using an emotional semantic context and whether this effect would be the same in the

NL and FL. The main task of the study required participants to listen to descriptions of countries and answer questions about them. Although participants performed better in their NL, results suggested that they benefited equally from the positive emotional context in both languages.

Preceding studies on the effects of emotionality on memory have mainly used visual stimuli. In contrast, the current study emulates information transfer in classroom settings by focusing on aural stimuli. Results showed statistically reliable emotionality effects with auditory information in both NL and FL. The partial eta squared of this effect is considered to be of medium effect size, within the context of educational research (Richardson, 2011). This corresponds to 10.9% of the variance explained and a practical difference of 3% on the current test. Although relatively discrete, this effect could be the difference between passing and failing an exam for a student that is struggling in a class. In more general terms, this study suggests that emotionally loaded semantic contexts—not just emotional content—conveying new pieces of information can improve memory.

Given that there are no studies addressing the particular questions of the current study—namely, looking at the effects of emotional context on content learning—the results need to be understood within the wider literature. The effects found here (NL: 2.7%, FL: 3.3%) were smaller than those of single-word studies with known words. In particular, these studies show effects between 7 and 26% in the NL and between 9.5 and 18% in the FL (Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Caldwell-Harris, 2009; Ferré et al., 2010)—with one exception showing a non-significant effect in the FL (Anooshian & Hertel, 1994). Studies manipulating emotional context rather than emotional content have found larger effects than the current one in recall (12%) but not in recognition—no accuracy difference, only in response time (Erk et al., 2003, 2005). On the other hand, studies on new word learning show smaller effects (2 – 3.5%), more similar to the ones in the current study (Ferré et al., 2015). Overall, these results suggest that the effects of emotionality are reduced when only the context is manipulated and when there is learning of new content, rather than repeating information that is already known. Therefore, our results are in accordance with those reported by prior literature and are within the predictable effect size.

The key result in this study is that the effect of emotionality is the same in the FL and the NL. This result is consistent with many recent studies using emotionality in single-word processing (Ayçiçeği & Harris, 2004; Caldwell-Harris, 2009; Ferré et al., 2015; Ponari et al., 2015), and suggest that this effect extends beyond individual word-learning to content learning. But, perhaps more importantly, this result challenges the view that the FL, in general terms, leads to emotional distancing (see Costa, Duñabeitia, & Keysar, 2018).

These results relate to the FLE and the theoretical issue of its origin. Hayakawa, et al. (2016) suggest that there are two main ways of explaining the FLE on moral decision-making: a reduction in emotional processing and increasing psychological distance. Both of these accounts would predict a reduced emotional effect in the FL compared to the native one. If emotionality is completely blocked, this described FLE would predict that emotionality and its effect on performance would be reduced or absent in the FL condition. With respect to psychological distance, the conclusion is the same: this would make the information seem more abstract, reducing the effect of emotionality. Therefore, neither of these ideas is consistent with our results—namely, an equal effect of emotionality in the NL and FL. On the other hand, if the FLE is circumscribed to only the manipulation of known information and its prior associations, it would explain why learning new information does not show the same effects. For example, learning the word “home” using neutral language would lead to more difficulty in learning it and a reduced emotional response for that word, whereas if it is presented using emotional language, perhaps it would be remembered better—showing an emotionality effect.

Looking at the results from this perspective, the current findings do not necessarily have to contradict the existence of the FLE. Instead, they suggest a possible mechanism for how it arises. Gawinkowska, et al.'s (2013) idea that the effect is due to social and cultural norm differences would suggest that emotionality should affect both language conditions equally in this case. This is consistent with our results, since if the FLE is circumscribed to differences in norms, it should not be present. Importantly, Geipel, et al.'s (2015a, 2015b, 2016) suggestion that the origin of this effect is a reduction of intuitive responses and a depletion of cognitive resources would imply a decrease in performance overall in the FL, but not necessarily any difference in emotionality. This reduction of cognitive resource availability explains our data better, predicting our decrease in performance in the FL, as well as the consistency of emotionality effects between languages.

In other words, the results of the current study could suggest that, rather than emotionality being reduced overall in an FL context, learners' cognitive resources are taxed, affecting emotionality differently according to the task. Furthermore, if the reduction in emotionality is observed in cases where only already-known information is concerned, perhaps it is because they are lacking emotional associations within that language. These results suggest that providing FL learners with more emotional materials—as in this case—could help them learn these associations.

It is worth noting that, although we did not intend to manipulate interest—and effectively the content was the same between conditions—perhaps the positive condition could have also presented the information in a more interesting way than the neutral one, contributing to the effect we found (see Hidi, 1990 for a review on the effect of interest on learning). In future studies, the effect of emotionality could be contrasted with that of “interest” or engagement. In addition, the effect we observe here might be increased further by engaging the participants in an activity where they have to use this new content or by making the information to be remembered self-relevant. For example, with the current materials, engagement could be increased by asking participants to not only listen passively but also to actively decide if they would want to move to the described country. Nevertheless, the current results open way for a new way of looking at both emotionality effects and learning in a foreign language which, with further replications, could provide a useful tool for teaching in a non-native language.

Conclusion

The current study reports a well-controlled experiment in line with CLIL approaches, as participants learned the same content in either their NL or an FL and were then tested using exactly the same task and materials. Learning in an FL may sometimes hinder memory of new content as a consequence of the difference in language knowledge and use with the NL. However, the use of emotional semantic contexts can be a short-term tool in the classroom, particularly during aural exercises or verbal transmission of new information in order to boost memory. Considering the emotional distancing or detachment that has been typically associated with FL contexts (see Costa et al., 2018), the use of emotionally loaded materials or activities in classroom settings could be useful for partially counteract existing FLEs.

Table 1: Matched Means and Standard Deviations

| | Self-Rated Level of English | | | | | | AOA of English | Spanish LexTALE | English LexTALE |
|------------------|-----------------------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|
| | Age | Listening | Reading | Speaking | Writing | Overall | | | |
| Foreign | 33.07 (8.91) | 7.07 (1.47) | 8.21 (0.93) | 6.81 (1.22) | 7.31 (1.18) | 7.15 (1.12) | 9.81 (3.77) | 0.94 (0.04) | 0.76 (0.08) |
| Native | 34.47 (9.63) | 7.31 (1.69) | 8.23 (1.26) | 7.05 (1.52) | 7.39 (1.53) | 7.39 (1.26) | 10.7 (6.71) | 0.93 (0.05) | 0.77 (0.09) |
| BF ₀₁ | 3.49 (0.01) | 3.50 (0.01) | 4.19 (0.01) | 3.30 (0.01) | 4.09 (0.01) | 3.05 (0.01) | 3.32 (0.01) | 3.20 (0.01) | 4.09 (0.01) |

Note: Numbers in parentheses refer to standard deviation for the FL and NL groups, except for in the final line (Bayes Factor) where they refer to error percentage. With BF₀₁ a positive number above 1 supports no difference between the two groups, with 3 and above implying moderate evidence that the means are equal. Age and age of acquisition of English are in years, the self-ratings of level of English are on a scale from 1 to 10, and the LexTALEs are scored from 0 (chance) to 1 (perfect score).

Table 2: Average Word Frequency by Language and Emotional Condition

| | Spanish | English |
|---|------------------|------------------|
| Neutral (M, SD) | 616.48 (1306.08) | 608.75 (847.98) |
| Positive (M, SD) | 727.03 (1793.12) | 641.29 (919.19) |
| Bayes Factor (BF ₀₁ , %error) | 7.29 (0.068) | 14.12 (8.63 e-6) |

Note: Numbers in parentheses refer to standard deviation for the FL and NL groups, except for in the final row (Bayes Factor) where they refer to error percentage. With BF₀₁ a positive number above 1 supports no difference between the two groups, with 3 and above implying moderate evidence that the means are equal.

Table 3: Average Accuracy in Percent Correct by Condition

| Language Condition | Emotionality | Mean | Standard Error | 95% Confidence Interval | |
|--------------------|--------------|-------|----------------|-------------------------|-------|
| | | | | Lower | Upper |
| English | Positive | 62.0% | 2.10% | 57.9% | 66.1% |
| | Neutral | 58.7% | 2.10% | 54.5% | 62.8% |
| | Overall | 60.3% | 1.90% | 56.5% | 64.2% |
| Spanish | Positive | 76.0% | 2.10% | 71.8% | 80.1% |
| | Neutral | 73.3% | 2.10% | 69.1% | 77.4% |
| | Overall | 74.6% | 1.90% | 70.7% | 78.5% |
| Total | Positive | 69.0% | 1.50% | 66.1% | 71.9% |
| | Neutral | 66.0% | 1.50% | 63.1% | 68.9% |
| | Overall | 67.5% | 1.61% | 64.3% | 70.6% |

References

- Admiraal, W., Westhoff, G., & De Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English 1. *Educational Research and Evaluation*, 12(1), 75–93. <https://doi.org/10.1080/13803610500392160>
- Aguilar, M., & Rodríguez, R. (2012). Lecturer and student perceptions on CLIL at a Spanish university. *International Journal of Bilingual Education and Bilingualism*, 15(2), 183–197. <https://doi.org/10.1080/13670050.2011.615906>
- Anghel, B., Cabrales, A., & Carro, J. M. (2016). Evaluating a bilingual education program in Spain: The impact beyond foreign language learning. *Economic Inquiry*, 54(2), 1202–1223. <https://doi.org/10.1111/ecin.12305>
- Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition & Emotion*, 8(6), 503–514. <https://doi.org/10.1080/02699939408408956>
- Ayçiçeği, A., & Harris, C. L. (2004). Bilinguals' recall and recognition of emotion words. *Cognition and Emotion*, 18(7), 977–987. <https://doi.org/10.1080/02699930341000301>
- Bergroth, M. (2006). Immersion students in the matriculation examination three years after immersion. In S. Björklund, K. Mard-Miettinen, M. Bergström, & M. Södergard (Eds.), *Exploring Dual-Focussed Education: Integrating language and content for individual and societal needs* (pp. 123–134). Vaasan yliopisto. Retrieved from http://www.uva.fi/materiaali/pdf/isbn_952-476-149-1.pdf
- Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*. <https://doi.org/10.1109/MIC.2008.114>
- Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism: Language and Cognition*, 12(3), 291–303. <https://doi.org/10.1017/S1366728909990125>
- Caldwell-Harris, C. L. (2014). Emotionality differences between a native and foreign language: Theoretical implications. *Frontiers in Psychology*, 5(1055), 1–4. <https://doi.org/10.3389/fpsyg.2014.01055>
- Costa, A., Duñabeitia, J. A., & Keysar, B. (2019). Language context and decision-making: Challenges and advances. *Quarterly Journal of Experimental Psychology*, 72(1), 1–2. <https://doi.org/https://doi.org/10.1177/1747021818789799>
- Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014). “Piensa” twice: On the foreign language effect in decision making. *Cognition*, 130(2), 236–254. <https://doi.org/10.1016/j.cognition.2013.11.010>
- Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *PLoS ONE*, 9(4), 1–7. <https://doi.org/10.1371/journal.pone.0094842>
- Costa, A., Vives, M.-L., & Corey, J. D. (2017). On language processing shaping decision making. *Current Directions in Psychological Science*, 26(2), 146–151. <https://doi.org/10.1177/0963721416680263>
- Dallinger, S., Jonkmann, K., Hollm, J., & Fiege, C. (2016). The effect of content and language

- integrated learning on students' English and history competences—Killing two birds with one stone? *Learning and Instruction*, 41, 23–31. <https://doi.org/10.1016/j.learninstruc.2015.09.003>
- Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning (CLIL) classrooms*. John Benjamins Publishing (Vol. 20).
- Dalton-Puffer, C. (2011). Content-and-language integrated learning: From practice to principles? *Annual Review of Applied Linguistics*, 31, 182–204. <https://doi.org/10.1017/S0267190511000092>
- Day, E., & Shapson, S. (1996). *Studies in immersion education (Vol. 11)*. Multilingual Matters.
- Dewaele, J.-M. (2010). *Emotions in Multiple Languages*. London: Palgrave Macmillan. <https://doi.org/10.1057/9780230289505>
- Erk, S., Kiefer, M., Grothe, J., Wunderlich, A. P., Spitzer, M., & Walter, H. (2003). Emotional context modulates subsequent memory effect. *NeuroImage*, 18(2), 439–447. [https://doi.org/10.1016/S1053-8119\(02\)00015-0](https://doi.org/10.1016/S1053-8119(02)00015-0)
- Erk, S., Martin, S., & Walter, H. (2005). Emotional context during encoding of neutral items modulates brain activation not only during encoding but also during recognition. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2005.02.045>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias Blanco, J. M. (2017). Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: empirical evidence from Spain. *International Journal of Bilingual Education and Bilingualism*, 0050, 1–14. <https://doi.org/10.1080/13670050.2017.1294142>
- Ferré, P., Garcia, T., Fraga, I., Sanchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition & Emotion*, 24(5), 760–785. <https://doi.org/10.1080/02699930902985779>
- Ferré, P., Ventura, D., Comesaña, M., & Fraga, I. (2015). The role of emotionality in the acquisition of new concrete and abstract words. *Frontiers in Psychology*, 6(976), 1–10. <https://doi.org/10.3389/fpsyg.2015.00976>
- Frances, C., de Bruin, A., & Duñabeitia, J. A. (n.d.). The effects of language and emotional context on vocabulary learning.
- Gawinkowska, M., Paradowski, M. B., & Bilewicz, M. (2013). Second language as an exemptor from sociocultural norms. Emotion-related language choice revisited. *PLoS ONE*, 8(12), 1–6. <https://doi.org/10.1371/journal.pone.0081225>
- Geipel, J., Hadjichristidis, C., & Surian, L. (2015a). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology*, 59, 8–17. <https://doi.org/10.1016/j.jesp.2015.02.001>
- Geipel, J., Hadjichristidis, C., & Surian, L. (2015b). The foreign language effect on moral judgment: The role of emotions and norms. *PLoS ONE*, 10(7), 1–17. <https://doi.org/10.1371/journal.pone.0131529>

- Geipel, J., Hadjichristidis, C., & Surian, L. (2016). Foreign language affects the contribution of intentions and outcomes to moral judgment. *Cognition*, *154*, 34–39. <https://doi.org/10.1016/j.cognition.2016.05.010>
- Hadjichristidis, C., Geipel, J., & Savadori, L. (2015). The effect of foreign language in judgments of risk and benefit: The role of affect. *Journal of Experimental Psychology: Applied*, *21*, 117–129. <https://doi.org/10.1037/xap0000044>
- Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (Vol. 56, pp. 257–283). Clevedon, UK: Multilingual Matters.
- Hayakawa, S., Costa, A., Foucart, A., & Keysar, B. (2016). Using a foreign language changes our choices. *Trends in Cognitive Sciences*, *20*(11), 791–793. <https://doi.org/10.1016/j.tics.2016.08.004>
- Hernandez-Nanclares, N., & Jimenez-Munoz, A. (2015). English as a medium of instruction: evidence for language and content targets in bilingual education in economics. *International Journal of Bilingual Education and Bilingualism*, *20*(7), 883–896. <https://doi.org/10.1080/13670050.2015.1125847>
- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, *60*(4), 549–571. <https://doi.org/10.3102/00346543060004549>
- Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica*, *35*, 49–66.
- Jäppinen, A.-K. (2005). Thinking and content learning of mathematics and science as cognitive development in content and language integrated learning (CLIL): Teaching through a foreign language in Finland. *Language and Education*, *19*(2), 148–169. <https://doi.org/10.1080/09500780508668671>
- Jeffreys, H. (1961). *Theory of Probability*, Oxford. Oxford, UK: Oxford University Press.
- Jiménez Catalán, R. M., & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL instruction. In *Content and language integrated learning: Evidence from research in Europe* (pp. 81–92).
- Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science*, *23*(6), 661–668. <https://doi.org/10.1177/0956797611432178>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, and Computers*, *28*(2), 203–208. <https://doi.org/10.3354/ame01683>
- Ouazizi, K. (2016). The effects of CLIL education on the subject matter (Mathematics) and the target language (English). *Latin American Journal of Content & Language Integrated Learning*, *9*(1), 110–137. <https://doi.org/10.5294/laclil.2016.9.1.5>

- Pavlenko, Aneta. (2002). Bilingualism and emotions. *Multilingua*, 21(1), 45–78.
- Pérez Cañado, M. L. (2018). The effects of CLIL on L1 and content learning: Updated empirical evidence from monolingual contexts. *Learning and Instruction*, 57, 18–33. <https://doi.org/10.1016/j.learninstruc.2017.12.002>
- Poirier, D. J. (2006). The growth of Bayesian methods in statistics and economics since 1970. *Bayesian Analysis*, 1(4), 969–980. <https://doi.org/10.1214/06-BA132>
- Ponari, M., Rodriguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2015). Processing advantage for emotional words in bilingual speakers. *Emotion*, 15(5), 644–652. <https://doi.org/10.1037/emo0000061>
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135–147. <https://doi.org/10.1016/j.edurev.2010.12.001>
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56(5), 356–374. <https://doi.org/10.1016/j.jmp.2012.08.001>
- Roussel, S., Joulia, D., Tricot, A., & Sweller, J. (2017). Learning subject content through a foreign language should not ignore human cognitive architecture: A cognitive load theory approach. *Learning and Instruction*, 52, 69–79. <https://doi.org/10.1016/j.learninstruc.2017.04.007>
- Schmitz, L. P. T. / C. (2019). LimeSurvey: An open source survey tool. Hamburg, Germany: LimeSurvey Project. Retrieved from <http://www.limesurvey.org>
- Sebastián-Gallés, N., Martí, M., Carreiras, M., & Cuetos, F. (2000). *LEXESP: Léxico Informatizado del Español*. Barcelona: Universitat de Barcelona.
- Serra, C. (2007). Assessing CLIL at primary school: A longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 10(5), 582–602. <https://doi.org/10.2167/beb461.0>
- Stohler, U. (2006). The acquisition of knowledge in bilingual learning: An empirical study on the role of language in content learning. *VIEWS-Vienna English Working Papers*, 15(3), 41–46. Retrieved from https://boris.unibe.ch/18531/1/views15_3_Stohler.pdf
- Surmont, J., Struys, E., Van Den Noort, M., & Van De Craen, P. (2016). The effects of CLIL on mathematical content learning: A longitudinal study. *Studies in Second Language Learning and Teaching*, 6(2), 319. <https://doi.org/10.14746/sslIt.2016.6.2.7>
- Van de Craen, P., Ceuleers, E., & Mondt, K. (2007). Cognitive development and bilingualism in primary schools: Teaching maths in a CLIL environment. In *Diverse Contexts—Converging Goals. CLIL in Europe* (pp. 185–200). Frankfurt am Main: Peter Lang.
- Vives, M.-L., Aparici, M., & Costa, A. (2018). The limits of the foreign language effect on decision-making: The case of the outcome bias and the representativeness heuristic. *PLoS ONE*, 13(9), 1–14. <https://doi.org/10.17605/osf.io/6mdvs.Funding>
- Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching*, 10(4), 116–126.
- Yang, S.-W. (2014). *The study of variables influencing the effect of English medium instruction on*

academic content learning and English proficiency development. The Pennsylvania State University.

Note: Participants showed no effect of order, $t(75) = .019$, $p = .891$, $BF_{01} = 7.85$, $error\% = 7.39 \times 10^{-6}$, showing moderate evidence that participants performed similarly regardless of order. Furthermore, there was moderate evidence that the two country descriptions were equally easy to remember, $t(75) = 1.23$, $p = .270$, $BF_{01} = 4.35$, $error\% = 5.15 \times 10^{-6}$.

Appendix: Number of emotional words and the average rating overall by language and condition

| | English | | | | Spanish | | | |
|----------|--------------------|-------------|--------------------|-------------|--------------------|-------------|--------------------|-------------|
| | High Valence Words | | High Arousal Words | | High Valence Words | | High Arousal Words | |
| | N | M(SD) | N | M(SD) | N | M(SD) | N | M(SD) |
| Neutral | 151 | 6.56 (1.03) | 81 | 5.01 (0.91) | 54 | 5.85 (1.71) | 37 | 5.14 (1.07) |
| Positive | 243 | 6.99 (1.01) | 176 | 5.43 (0.95) | 193 | 7.21 (1.06) | 172 | 6.07 (1.08) |

Note: N stands for the number of words with values >5. The means and standard deviations are overall on a scale from 1 to 9.