Self-bias and the emotionality of foreign languages

Lela Ivaz[1*]

Kim L. Griffin[2]

Jon Andoni Duñabeitia[1,3]

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

[2]Universidad Europea del Atlántico (UNEATLÁNTICO); Santander, Spain

[3] Facultad de Lenguas y Educación, Universidad Nebrija; Madrid, Spain

*Contact information:

Lela Ivaz

l.ivaz@bcbl.eu

Basque Center on Cognition, Brain and Language (BCBL)

Paseo Mikeletegi 69, 2nd floor

20009 Donostia - Spain

+34 943309300 (ext. 321)

Abstract

Foreign language contexts impose a relative psychological and emotional distance in bilinguals. In our previous studies, we demonstrated that the use of a foreign language changes the strength of the seemingly automatic emotional responses in the self-paradigm, showing a robust asymmetry in the self-bias effect in a native and a foreign language context. Namely, larger effects were found in the native language, suggesting an emotional blunting in the foreign language context. In the present study, we investigated the source of these effects by directly comparing whether they stem from a language's foreignness *vs.* its non-nativeness. We employed the same self-paradigm (a simple perceptual matching task of associating simple geometric shapes with the labels "you," "friend", and "other"), testing unbalanced Spanish-Basque-English trilinguals. We applied the paradigm to 3 language contexts: *native*, non-native but contextually-present (i.e., *non-native local*), and *non-native foreign*. Results showed a smaller self-bias only in the foreign language, pointing to the foreign-language-induced psychological/emotional distance as the necessary prerequisite for foreign language effects. Furthermore, we explored whether perceived emotional distance towards foreign languages in Spanish-English bilinguals modulates foreign language effects. Results suggest that none of the different indices of emotional distance towards the foreign language obtained via questionnaires modulated the self-biases in the foreign language contexts. Our results further elucidate the deeply-rooted and automatic nature of foreign-language-driven differential emotional processing.

Keywords: foreign language effects, self-bias, non-native languages, emotionality

Word count: 8,134 (text) + 1,407 (references)

Acknowledgements

Introduction

Most bilinguals experience less discomfort than native speakers when discussing emotionally-distressing, embarrassing or traumatic experiences, as well as when uttering or hearing demeaning remarks (e.g., swear words, taboo words, insults, childhood reprimands, etc.) in a language that is not native to them (e.g., Bond & Lai, 1986; Dewaele, 2004a, 2004b; Dewaele & Pavlenko, 2002; Pavlenko, 2006, 2007). In this case, foreign languages serve as an emotional buffer that absorbs the stark emotional impact of affective states, making bilinguals feel less uneasy when dealing with highly emotionally-charged language (see Iacozza, Costa, & Duñabeitia, 2017, for review). Observations and reports of this behavior over the years have sparked a keen interest among the scientific community in the subject of decreased emotional affect and foreign language contexts.

Foreign language effects. This phenomenon of attenuated emotional reactivity in bilinguals when they are in a foreign language context, known as foreign language effect, is a topic on which there is a rapidly growing literature. What started out as anecdotal notes on bilinguals feeling emotions less keenly and being less emotionally affected overall in their foreign languages, has been experimentally corroborated with a series of experiments covering a range of phenomena, including the perception of emotive language such as swear words and taboo words (Caldwell-Harris, 2014; Harris, Ayçiçegi, & Gleason, 2003; for a review, see Pavlenko, 2012). Also, bilinguals have reported how profoundly different it feels to take part in emotionally-charged behaviors (e.g., declaring love to someone, praying, or lying) in their native language (Caldwell-Harris, 2014; Dewaele, 2010; Pavlenko, 2005), and the impact of foreign language effects is also being explored using psychophysiological measurements (e.g., skin conductance or pupil dilation, such as in García-Palacios et al., 2018; see also Harris, 2004; Harris et al., 2003).

The scope of the foreign language effects even extends to moral decisions (e.g., Corey et al., 2017; Costa, Foucart, Arnon, et. al., 2014; Costa, Foucart, Hayakawa, et al., 2014; Geipel, Hadjichristidis, &

Surian, 2015a, 2015b; Keysar, Hayakawa, & An, 2012; Shin & Kim, 2017). These studies speak volumes to the remarkable pervasiveness and comprehensiveness of the foreign language effects. Namely, the effects have been shown to be capable of affecting even our moral principles, which are assumed to occur at a much more fundamental, conceptual level than the one governing language processing. Intuitively, it would stand to reason to think that our moral decisions are completely divorced from language processing in every conceivable way – yet, all of the evidence in the field seems to point to the contrary. When moral dilemmas are couched in a foreign language, irrational and impulsive decisions are greatly reduced (e.g., Corey et al., 2017). In fact, the mere presence of foreign language contexts can obliterate several cognitive framing biases (e.g., Keysar, Hayakawa, & An, 2012). Interestingly, the differences between languages in experiments on moral dilemmas become less pronounced as foreign language proficiency increases (Costa, Foucart, Hayakawa et al., 2014; Geipel et al., 2015b).

Suggested explanations to foreign language effects. We usually speak and comprehend foreign languages less fluently and more strenuously than our native ones, which is why the process has usually been characterized as cognitively draining. Namely, foreign language processing incurs an increase in the processing disfluency, prompting more deliberate cogitation, and as such is conceived as depleting our cognitive resources (e.g., Green, 1998; Hernandez & Meschyan, 2006; see Duñabeitia & Costa, 2015, for several markers of the cognitive load imposed by foreign language speech). In terms of the Dual-Process Theory (e.g., Kahneman, 2003, 2011; Sloman, 1996), foreign languages seem to engage more the analytical, effortful, rule-governed, and cogitation-heavy processing routines (i.e., System 2; Costa, Foucart, Arnon, et. al., 2014), as opposed to the reflexive, effortless, emotion-laden and intuition-governed cognitive method (i.e., System 1). Hence, foreign language processing prompts a great deal of deliberation.

But what is the link between the increased cognitive strain and emotional detachment? System 2

entails the more rational and dispassionate cognitive route, and its activation in conjunction with foreign language processing was initially thought to be the ultimate driving force mediating the dampened emotional resonance in foreign languages. However, in light of recent evidence, it seems that it is System 1 that is behind the emotional blunting in foreign language. Specifically, it has recently been shown that foreign language use does not boost System 2 as much as it hampers System 1 processing. Foreign language contexts hinder emotional processing and deontological considerations intrinsic to System 1 (Hayakawa et al., 2017), thus giving rise to the foreign language effects.

The literature on foreign language effects has yet to reach a definitive consensus on the underlying source of emotional blunting. Most authors subscribe to the predominant theory hinting at the emotionally-poor, structured, and impersonal academic settings in which foreign languages are acquired, as the origin of the reduced emotionality in the foreign language context (e.g., Caldwell-Harris, 2014; Dewaele, 2004b, 2010, 2013). In fact, this learning context is the most important difference between a foreign language and a non-native (i.e., second) language, given that the former is typically learned in academic settings which reduce the chances of its use in everyday situations (e.g., Caldwell-Harris, 2014; Dewaele, 2010; Ivaz, Costa & Duñabeitia, 2016; see also Duñabeitia, 2017). The familial context in which native language acquisition is embedded provides the backdrop that runs the gamut of all human emotions in an infant's emotionally formative early years. The simultaneous development of the emotional system and linguistic skills makes native language and emotions inextricably interwoven. This is probably why foreign languages are not as rooted in the emotional system.

Hence, it stands to reason that the emotional resonance in one's native language would be in sharp contrast to the one in foreign languages. To this point, Ivaz, Costa & Duñabeitia (2016) argue that foreign language-related highly limited contextual scope is at the core of the emotional blunting in

the foreign language effects. Note, though, that this is not to say that a language cannot elicit high

emotional reactivity much later in life and beyond the most emotionally formative years in the early

childhood (Caldwell-Harris, 2008; Pavlenko, 2005). Along similar lines, some authors argue that the

frequency of foreign language use plays an important part in its emotional resonance (e.g., Degner,

Doycheva, & Wentura, 2011; Puntoni, de Langhe, & van Osselaer, 2009), while we argued that the

lack of contextual diversity in its use needs to be singled out as a significant contributor to the

reduced emotionality it tends to spawn (e.g., Ivaz, Costa & Duñabeitia, 2016). If people do not tend

to use their foreign language in everyday situations, its use is less likely to favor emotionality.

Foreignness vs. non-nativeness. Previous work in the field notwithstanding, much is still unclear

about what ultimately drives the foreign language effects. Some studies have demonstrated that the

higher the proficiency in one's foreign or second language, the more similar the emotional response

in these languages and the native one (Harris, Gleason, & Ayçiçegi, 2006; in studies on moral

dilemmas, Costa, Foucart, Hayakawa et al., 2014; Geipel et al., 2015b). For instance, it is still

unknown whether it is the **foreignness** or the **non-nativeness** of foreign languages that drives

these effects and elicits the reduced emotionality observed therein. In this case, foreignness refers

to the relative psychological distance towards the foreign language in question and the culture it

represents, whereas non-nativeness indicates acquiring a language through immersion and

sequentiality, as opposed to the same way as one's native language is acquired – from birth. Foreign

and second languages could feasibly yield similar effects, which is an interesting question that

merits more attention in the research on foreign languages. This outcome of similar effects in

particular would privilege the non-nativeness hypothesis, providing support for the claim that

acquiring a language after early childhood may be the source of the blunted affect tied to foreign

languages. The premise of pitting foreign and second language contexts against each other could

help us to draw a clearer demarcation line between the languages and help us to understand the

origin of muted emotionality better. The current study builds on this question by using a paradigm

that has been shown to elicit robust foreign language effects: the self-paradigm.

The self-paradigm. Self-image and self-centeredness are phenomena whose underlying neural mechanisms overlap with those responsible for emotional responses. These phenomena are special in that any stimuli pertaining to them consistently elicit overwhelmingly favorable responses to them. The literature abounds with examples of this self- bias effect. Namely, in a series of recent experiments, a simple geometric shape-label association paradigm was used to show a pronounced prioritization in the processing of stimuli when they are self-related as opposed to related to other people (Sui, & Humphreys, 2015a, 2015b; Sui, He & Humphreys, 2012; Sui, Rotshtein, & Humphreys, 2013; Sui et al., 2014, 2015). The authors randomly assigned neutral geometric shapes (i.e., triangles, circles, and squares) to the participant and two other people: their friend/mother and a stranger. Participants were taught the correct matchups and subsequently shown random combinations of shapes and labels so as to indicate whether they matched or mismatched the instruction. The rationale behind the labels was to include a person with whom participants have a close relationship (i.e., their mother or their friend), and another random person to whom participants would not have any type of connection.

This design consistently gives rise to the so-called self-bias effect, an outcome where the fastest and most accurate responses by a large margin are those to self-related stimuli in matching trials. The effect is so robust that it holds even when the relative number of these stimuli is decimated (Sui et al., 2014), speaking to the sheer automaticity of the effect. Moreover, its magnitude grows when the size of the stimuli is increased (Sui, & Humphreys, 2015a). The self-bias effect is akin to the effect elicited by high-value-reward labels (e.g., triangle 8£ vs. square 1£; Sui, He, & Humphreys, 2012), and occurs even with other kinds of stimuli (e.g., simple movements related to self; Frings, & Wentura, 2014). The self-bias paradigm demonstrated how associating simple visual stimuli with the self, drastically increases their social (and perceptual) salience, thus boosting memory, speed,

and accuracy in participants.

What is it that would make the self-paradigm suitable for exploring the foreign language effects? In different areas of psychology research, the concept of self-esteem and self-image has been found to arouse emotions that run the gamut of emotions from happiness, pride, and content, to contempt, hostility, depression, and anxiety (e.g., Bushman & Baumeister, 1998; Diener & Diener, 1995; Kernis, Grannemann, & Barclay, 1989; Mineka, Watson, & Clark, 1998; Tangney & Fischer, 1995). The self-bias seems to tap into our emotionality, since it can be conceived of as self-emotions. Specifically, our systematically biased responses to the information related to us have been interpreted as reactions to intrinsically rewarding information (Northoff & Hayes, 2011), a hypothesis attested to by Sui, He, & Humphreys (2012) study. Moreover, the responses that self-related stimuli elicit are indeed similar to those brought forth by reward-related stimuli (Sui, & Humphreys, 2015b). There also seems to be neural evidence offering support to the emotional component of the self-bias, since the areas of the brain that mediate self-bias partially overlap with those specializing in emotional reactions (e.g., the prefrontal cortex; Sui, Rotshtein, & Humphreys, 2013).

The present study. In our previous studies, we were interested in seeing whether this self-bias would be modulated by the foreign language context. If this context triggers psychological and emotional distance in bilinguals, and if the self-bias is governed by our emotional system, then there should be a systematic difference in the self-biases in the native and foreign language contexts. To this end, we conducted several between- and within-subject experiments where we looked at the self-bias magnitudes in the two languages (Ivaz, Costa & Duñabeitia, 2016). The size of the self-bias was halved in the foreign language context in the between-subject design, while it was still significantly diminished even in the within-subject design. Similar results were replicated in a more recent study by Shin & Kim (2017) with Korean-English bilinguals. However, our previous

study was not able to fully distinguish between the effects of foreign vs. non-native languages, since our participants' non-native language was also a foreign language (i.e., English as a non-native language of native Spanish speakers).

The focal question of Experiment 1 is whether the root of these previously reported effects is a language's foreignness or non-nativeness. In the current study, we tried to shed some light on this issue, and to do so, in Experiment 1 we tested three different types of languages: native, second (i.e., non-native local), and foreign (i.e., non-native foreign). We used the same paradigm as we did in Ivaz, Costa, & Duñabeitia (2016) and applied it to three language contexts: native language (Spanish), non-native but contextually present language to which there is a daily exposure and an emotional attachment – non-native local (Basque), and foreign language – non-native foreign (English). The aim of Experiment 2 was tackling the notion of perceived emotional distance towards foreign languages. We were interested in seeing whether a more positive attitude towards one's foreign language could modulate the magnitude of the foreign language effect. We employed the same self-paradigm and compared the effects in the native and foreign languages as a function of the emotional attachment to each of them as reported by participants in a series of questionnaires.

Experiment 1

The main questions in this experiment referred to whether the self-bias was a case of foreign or non-native language use, and whether it was modulated by the social context or not. To this end, we tested native Spanish speakers from the Basque Country who also knew Basque and English. The Basque Country (i.e., the Basque Autonomous Region) in the north of Spain has two co-official languages: Spanish and Basque. Both languages are equally represented across the community: all signs are bilingual, there are bilingual schools, all public service employees are bilingual, etc. Despite that, it can be said that Spanish is the dominant language in several aspects. According to

the demographic statistics published by the Basque Government in 2014, Spanish remains the predominant language of the region. Namely, it is understood by virtually 100% of the inhabitants, 72.6% have it as their native language and 77.1% report speaking it at home. On the other hand, Basque is understood by 55.7% of the population, it is the native language of 23.5%, and 20.8% speak it at home (The Basque Government, 2014). All of our participants had Basque as a non-native second language and Spanish as their mother tongue. It stands to reason that our participants would foster an emotional attachment towards Basque, to a varying degree. Intuitively, we would expect there to be a clear difference between two non-native languages learned in strikingly different settings: natural (i.e., the immediate environment of a contextually present non-native language) as opposed to artificial (i.e., the classroom environment of a foreign language). Being in contact with a language on a daily basis and in varying circumstances of social interaction would mimic the contexts in which native language is normally used and acquired, tapping into the same well of emotional investment and emotionally-charged interactions. This would equate the second language to the native one, and put them both in stark contrast to the foreign language, which is generally used in specific, fairly limited, and academic situations. This was indeed the case of English for the participants tested in the current experiment, given that they had acquired English as a foreign language.

There are two possible sources of the foreign language effect that the self-bias paradigm might yield. Specifically, it is unclear whether the effect is driven by the foreignness of the language contrasted with the native one, or by the non-nativeness of that language. By testing how the use of a native vs. a second vs. a foreign language modulates the self-bias, we will be able to tell if the effect truly is a foreign language effect or a non-native language effect instead. If the source of this effect is the foreignness of a language, there will be little difference between the effect magnitudes in the native and the second language. If, on the other hand, it is a matter of non-nativeness, then the size of the non-native effect obtained will, in all likelihood, be equal to or smaller than the

foreign language effect.

Method

Participants. Twenty-eight participants[1] (16 females, mean age=24.04, SD=4.87) with normal or corrected-to-normal vision took part in the study. They were all non-balanced Spanish-Basque-English trilinguals. They were native Spanish speakers born and raised in the Basque Country, who were also relatively highly (and approximately equally) proficient in Basque and English. Prior to administering the experimental task to our participants, we inquired into their linguistic profile by assessing their proficiency in the three languages with different measurements (see Table 1; cf. Procedure in Ivaz, Costa, & Duñabeitia, 2016). Among these measures were their age of acquisition of Basque and English (in years, which were significantly different, as could be expected from participants living in a bilingual community), vocabulary knowledge as assessed by their picture naming scores (i.e., the so-called BEST test, a vocabulary test that included a battery of 65 drawings of common concepts, cf. de Bruin, Carreiras, & Duñabeitia, 2017; Gollan et al., 2012) with no significant differences between Basque and English, individual interviews' scores (as assessed by a bilingual interviewer), self-reported scores of their overall proficiency (in reading, writing and speaking) in the languages on a 10-point Likert-like scale, and self-reported daily exposure to the languages (expressed in % of time), among others. All of our participants completed the task in three languages.

- Insert Table 1 here -

Materials and Procedure. The materials employed in this within-subjects experiment were very similar to those we used in the Experiment 2 in Ivaz, Costa, & Duñabeitia (2016). Our materials

---

[1] In our previous paper (Ivaz, Costa, & Duñabeitia, 2016), we had 48 participants in the within-subject Experiment 2. While it would have certainly been preferable to have as many participants in the current experiments as in the aforementioned study. However, due to very stringent participant inclusion criteria, we ended up with a lower number. Nonetheless, as shown in the right panel of the Figure 2 in Ivaz Costa, & Duñabeitia (2016), the magnitude of the self-bias effect is relatively consistent across participants, suggesting that the magnitude of the samples may not be a critical factor.

were simple, white, unfilled geometric shapes (i.e., triangles, squares and circles) whose dimensions were 263 x 263 pixels, a fixation point (i.e., the "+" sign) located below the shapes, and written labels (in English "you," "friend," "other"; in Basque "zu," "laguna," "bestea"; in Spanish "tú," "amigo," "otro") in white, lowercase, Courier New font displayed against a black background (see Figure 1). In an effort to simplify the procedure, we opted for only one experimental list. All our participants completed a total of 540 trials, 180 per each of the three languages. Out of these 180 trials, 90 were matching (i.e., trials whose shape-label combinations matched the instruction) and 90 mismatching (i.e., random combinations of shapes and labels not corresponding to the instruction). The order of trials and languages was completely randomized.

-Insert Figure 1 here-

The data collection was run on Experiment Builder, v. 1.10.1630 (SR Research, Mississauga, ON, Canada) on a 19" CRT screen (with 1024 x 768 resolution at 100 Hz frequency), and with a response box. The instructions were presented in all three languages. Participants were told to remember the correct shape-label combinations, and imagine that they, themselves, are represented by a square, that their best friend is a circle, and that someone they do not know is a triangle. When presented with a combination, they were told to indicate whether it matched or mismatched the instruction as fast and as best as they could. Then, they completed a brief practice session of 36 trials, and in case they had more than 30% of errors, they had to redo the session. Each trial would start with the fixation point presented for 500 ms. This was followed by a geometric shape, the fixation point, and a label appearing at the same time for 1200-1500 ms. Feedback followed and stayed on screen for 500 ms (see Figure 2). Depending on the response, it consisted of either a happy (for correct) or a sad emoji (for incorrect responses or timeouts). The timeouts were randomized and set between 1200 and 1500 ms. Participants had 8 small breaks in total, one following every 60 trials. The experiment took approximately 20 minutes to complete.

-Insert Figure 2 here-

Results

We hypothesized that, unlike language non-nativeness, language foreignness would reduce the self-bias in our participants. Hence, we expected a significant difference in the size of the self-bias effects for the native and non-native local languages when compared to the non-native foreign language. ANOVAs were run on the latency and accuracy data for the matching trials following a 3 (Shape Category: self/familiar/other) × 3 (Language Context: native/second/foreign) design. Pairwise comparisons were also run on the data corresponding to the mismatching trials collapsed across conditions to determine the presence or absence of a language effect. RTs associated with erroneous responses and those above or below the cut-off points (mean ± 2.5 SD in each condition for each participant), were removed from the latency analysis (1.63% of outliers). Group-based means per condition are reported in Table 2.

- Insert Table 2 here -

Matching trials. The general 3×3 ANOVA on the RTs showed a main effect of Shape Category, $F(2,54)=86.77$, $p<.001$, partial $\eta^2=.763$, $1-\beta=1$. The main effect of Language Context was not significant, $F(2,54)=1.69$, $p=.19$. The interaction between these two factors was significant, $F(4,108)=3.83$, $p<.05$, partial $\eta^2=.12$, $1-\beta=.88$, revealing that the magnitude of the effects varied as a function of the language (native vs. second vs. foreign). A 2×3 ANOVA performed to explore _self-bias effects_ (Shape Category: self/other; Language Context: native/second/foreign) showed that participants responded faster in the self than in the other condition, $F(1,27)=133.63$ $p<.0001$, partial $\eta^2=.83$, $1-\beta=1$. There was a significant interaction between the two factors, $F(2,54)=6.9$, $p<.0005$, partial $\eta^2=.2$, $1-\beta=.91$. Additional pairwise comparisons were performed to pinpoint the loci of the significant interaction. All the pairwise comparisons were Bonferroni-corrected. The size of the self-effect was virtually identical in the native- (110 ms) and second-language context (111

ms), t(27)=-.13, p=.9. The self-effect was significantly larger in the native- (110 ms) than in the foreign-language condition (77 ms), t(27)=3.37, p<.005, as was the case in the second- (111 ms) vs. the foreign-language (77 ms) condition, t(27)=3.43, p<.005. A 2×3 ANOVA performed to explore *familiar-bias effects* (Shape Category: familiar/other; Language Context: native/second/foreign) showed a main effect of shape, indicating that participants were also significantly faster in the familiar than in the other condition, F(1,27)=21.78, p<.0001, partial η^2=.45, 1–β=.99. The main effect of language was not significant, F(2,54)=.032, p=.97. The interaction between these two factors was significant, F(2,54)=4.05, p<.05, partial η^2=.13, 1–β=.7, revealing that the magnitude of the effects varied as a function of the language (native vs. second vs. foreign). Additional pairwise comparisons were performed to find the sources of the significant interaction. All the pairwise comparisons were Bonferroni-corrected. The size of the familiar-effect was similar in the native- (35 ms) and second-language interaction (45 ms), t(27)=-.82, p=.42. The familiar-effect was significantly larger in the native- (35 ms) than in the foreign-language interaction (14 ms), t(27)=2.08, p=.047, as well as in the second- (45 ms) as compared to the foreign-language interaction (14 ms), t(27)=2.94, p<.05.

The analysis on the error rates following the general 3×3 design showed a significant main effect of Shape Category, F(2,54)=19.68, p<.0001, partial η^2=.42, 1–β=.99, and no effect of Language Context, F(2,54)=.36, p=.62. There was a significant interaction between these two factors, F(4,108)=2.98, p<.05, partial η^2=.1, 1–β=.78. A follow-up 2×3 ANOVA showed that participants responded more accurately in the *self* (5.48%) than in the *other* condition (15.52%), F(1,27)=31.54, p<.0001, partial η^2=.54, 1–β=.1, while there was no main effect of Language Context, F(1.64, 44.26)=2.37, p=.11. This *self-bias effect* was significantly different only in the native- vs. foreign-language trials comparison (12.02% vs. 7.02%, respectively), t(27)=2.46, p<.05. There was a marginally significant interaction between these factors, F(2,54)=3.14, p=.051, partial η^2=.1, 1–β=.58. Additional, Bonferroni-corrected pairwise comparisons were performed to explore this interaction. The error-rate self-

bias was similar in the native- (12.02%) and second-language context (11.07%), t(27)=-.47, p=.64. The error-rate self-bias was significantly larger in the native- (12.02%) than in the foreign-language condition (7.02%), t(27)=2.46, p<.05, but the same did not occur in the second- (11.07%) vs. foreign-language (7.02%) condition, t(27)=1.76, p=.09. Also, participants responded more accurately to the stimuli in the *familiar* than in the *other* condition, F(1,27)=9.7, p<.005, partial η^2=.26, 1–β=.85, without a main effect of Language Context, F(2,54)=.68, p=.51. There was a significant interaction between the two factors, F(2,54)=4.47, p<.05, partial η^2=.14, 1–β=.74. Additional, Bonferroni-adjusted, pairwise comparisons of the *familiar-bias effects* showed that our participants made significantly more errors in the familiar condition in their native (7.38%) and second (8.33%) languages, rather than in their foreign language (1.9%), t(27)=2.45, p<.05, and t(27)=2.86, p<.01, respectively.

Mismatching trials. Mismatching trials were responded to slightly more quickly in the native than in the second (t(27)=-3.3, p<.005) or foreign language (t(27)=.033), 752 ms vs. 765 ms vs. 762 ms; F(2,54)=4.8, p<.05, partial η^2=.15, 1–β=.77. There were significant differences between languages observed in the accuracy data (10.79% vs. 12.7% vs. 10.99% of errors, respectively; F(2,54)=3.41, p<.05, partial η^2=.11, 1–β=.62. Significantly fewer errors were made in the native (t(27)=-2.6, p<.05) and foreign (t(27)=2.26, p<.05) languages, as opposed to the second language condition.

<div align="center">

Discussion

</div>

The results of this experiment showed a sizeable reduction of the self-bias in the non-native foreign but not in the non-native local language. Self-bias is thought of as a product of the foreign language effect, and these results show that it does not extend to second languages. It is true that neither our second nor foreign languages are our native languages. However, it seems that the foreign-language-related physical distance from an everyday, organic, linguistic context seems to connote an emotional distance towards that language as well. The effect is the same in the two contextually

present languages, and the only difference observed was in the contextually absent language – in this case, the non-native foreign language.

In their skin conductance study, Harris, Gleason, & Ayçiçegi (2006) found virtually no difference in the way early bilinguals processed emotional phrases in their native language and their non-native but contextually-present second language. Our results are very much in line both with this study and the observations that the most striking differences in emotionality are between the native and foreign languages. Moreover, language immersion and frequent, contextually diverse language use can engender strong emotional attachment to languages that are not our native (e.g., Caldwell-Harris, 2014; Degner, Doycheva, & Wentura, 2011; Dewaele, 2010). This frequent and diverse language use prompts a shift in the way we encode our concepts, making those belonging to our native language overlap with those of our second one. This in turn causes both our native and non-native languages to elicit highly similar kinds of emotional response (Pavlenko, 2005, 2006). Furthermore, these results also speak to high proficiency and frequent language use being the key prerequisites for increased emotionality in a non-native language.

Experiment 2

The results from Experiment 1 demonstrated that the self-bias difference was unequivocally driven by the foreign language effect. Following up on these results, we wanted to investigate the interplay between the magnitude of this effect (as observed in the self-bias) and the language attitude, conceived as the attitude bilinguals develop towards a given language they speak (Dewaele, 2009). We hypothesized that the magnitude of the self-bias effect would vary as a function of the language attitude in an inversely proportional manner – the more positive (i.e., highly scored) the language attitude, the smaller the foreign language effect should be. In other words, the more emotionally affected our participants are in their foreign language, the more similar the effect elicited by this language will be to the native language.

Method

Participants. Sixty-four non-balanced Spanish-English bilingual students (47 females) of the Universidad Europea del Atlántico participated in the present study (mean age M=19.14 years, SD=1.75). They all had normal or corrected-to-normal vision, a relatively high proficiency of English and spoke it as a foreign language. We provide more information about the linguistic profile of our participants in Table 1. None of the participants had spent more than 6 months in an English-speaking country, since we suspected that this type of experience could induce a stronger emotional attachment to the foreign language. Specifically, it could do so by virtue of changing the status of English as a foreign language into a second language due to immersion and everyday exposure to it. Also, we only included the participants whose general English proficiency ranged from 4 to 8 according to a self-reported proficiency measure (see Materials and Procedure). We thus avoided the inclusion of participants with native-like performance, and more importantly, those with such a low proficiency that would be detrimental to the comprehension of the instructions.

Materials and Procedure. Prior to taking part in the self-bias experiment in Spanish (native language) and English (foreign language) similar to that reported in Ivaz et al.'s (2016) Experiment 2, our participants filled out a questionnaire in Spanish about their linguistic habits, their emotional attachment to the native and to the foreign language, and their empathy level in general. The questionnaire we used was an adapted version of Dewaele & Pavlenko's (2001-2003) Bilingualism and Emotions questionnaire that also included Baron-Cohen & Wheelwright's (2004) Empathy Quotient test. Much like its predecessors, our version of the questionnaire contained a number of open- and closed-ended questions with Likert-scale-like answers aimed at specifying one's language attitudes or the degree of empathy they feel in different situations (see Appendix).

Some additional sociodemographic questions were also added to gather as much data on our

participants as possible. Having our participants complete the questionnaire in Spanish allowed us to circumvent the possibility of our respondents rating their emotions more intensely in their non-native language – a systematic bias that has been known to occur in bilinguals' non-native languages (i.e., anchor contraction effect, De Langhe et. al., 2011; Harzing, 2006).

We obtained 3 types of self-reported variables assessing participants' socio-demographic and linguistic background. These variables were: **linguistic** (i.e., the *age of acquisition of English* in years (M=4.72, SD=1.61), the *general English proficiency* according to a 10-point Likert scale (M=6.12, SD=1.36), and the *exposure to English* - the percentage of time of daily exposure to it (M=26.34, SD=19.41)), **affective** (i.e., the *empathy quotient* (M=42.67, SD=9.42)), and **affective-linguistic** (i.e., the *language-based emotional reactivity score* (M=0.66, SD=0.25), the *self-perceived emotionality of English* (M=3.25, SD=1.21), and the *impact of emotional expressions in English* (M=2.68, SD=0.82)) (see Appendix for details). The *language-based emotional reactivity score* was calculated based on the responses to a group of 10 questions that asked respondents to opt for one of the possible languages (Spanish, English or other) in certain communicative scenarios. These questions asked them to select the language in which they preferred to express certain emotions. Spanish was scored 1, English -1 and any other language was scored 0. This way, the average score determined the language-based emotional reactivity score of a given participant, such that the more positive the score, the more emotionally bound to Spanish was the participant. A negative score indicated an emotional attachment to English. The *self-perceived emotionality of English* variable corresponded to the score respondents gave to this question: "*Indicate how much you agree with the following statement: English is emotional".* The possible answers to this question also ranged from 1 (none) to 5 (extreme). The *impact of emotional expressions in English* was calculated based on a group of nine questions which explored emotional language in English. The possible answers ranged from 1 (none) to 5 (extreme).

The questionnaire was administered to the participants at least 48 hours before they completed the experiment with the self-paradigm. This was done to avoid any carryover effects some questions could potentially have on our participants' performance in the self-bias task. The critical part of the experiment with the self-bias paradigm took approximately 12 minutes to complete. The materials and the procedure we employed were the same as in Experiment 1 (see Figures 1 and 2), albeit with only two language conditions: native vs. foreign (i.e., Spanish and English words).

Results

A two-step analysis routine was followed. First, we once again replicated the difference in the self-bias effect in native vs. foreign languages by analyzing the latency and accuracy data belonging to the critical matching trials, as in our preceding experiments. Second, a linear regression was performed to ascertain whether the 7 variables derived from the questionnaires were successful predictors of the self-bias effects in the foreign language. If the affective-linguistic factors related to the foreign language in any manner determine or modulate the self-bias effect, we would expect a difference in its magnitude. Namely, participants with a more positive attitude towards English should display a larger self-bias effect in English as compared to the ones with a more positive attitude towards Spanish.

RTs associated with erroneous responses and those above or below the cut-off points (mean ± 2.5 SD in each condition for each participant) were removed from the latency analysis (1.37% vs. 1.01% of outliers in the native- and foreign-language contexts, respectively). Group-based means per condition are reported in Table 2.

First, a series of ANOVAs were performed following a 3×2 design: 3 (Shape Category: self/familiar/other) × 2 (Language Context: native/foreign). The general 3×2 ANOVA on the RTs revealed a main effect of Shape Category, $F(1.83, 115.04)=114.41$, $p<.0001$, partial $\eta^2=.69$, $1-\beta=1$. The main effect of Language Context was not significant, $F(1, 63)=2.28$, $p=.14$. The two-way

interaction was statistically significant, $F(1.76, 110.44)=4.31$, p<.05, partial $\eta^2=.06$, $1-\beta=.7$. A 2×2 ANOVA performed to explore self-bias effects (Shape Category: self/other; Language Context: native/foreign) showed that participants responded faster in the self than in the other condition, as indicated by the main Shape Category effect, $F(1, 63)=205.03$, p<.0001, partial $\eta^2=.76$, $1-\beta=1$. There was no main effect of Language Context, $F(1, 63)=3.23$, p=.08. There was, however, a significant Shape Category × Language Context interaction, $F(1, 63)=6.66$, p<.05, partial $\eta^2=.1$, $1-\beta=.719$. This suggested that the self-bias effect was larger in the native than in the foreign language (134 ms vs. 113 ms; $t(63)=2.58$, p<.05). A 2×2 ANOVA performed to explore familiar-bias effects (Shape Category: familiar/other; Language Context: native/foreign) showed that participants responded faster in the familiar than in the other condition, $F(1, 63)=14.13$, p<.0001, partial $\eta^2=.18$, $1-\beta=.96$, while no difference was observed between the two language contexts, $F(1, 63)=.12$, p=.73. No interaction was observed between the groups, $F(1, 63)=.49$, p=.49.

The analysis on the error rates following the general 3×2 design showed a significant main effect of Shape Category, $F(1.8, 113.46)=64.97$, p<.0001, partial $\eta^2=.51$, $1-\beta=1$, but there was no effect of Language Context, $F(1, 63)=2.4$, p=.13. However, there was a significant interaction between these two factors, $F(2, 126)=5.8$, p<.005, partial $\eta^2=.08$, $1-\beta=.86$. A follow-up 2×2 ANOVA showed that participants responded more accurately in the *self* than in the *other* condition, $F(1, 63)=105.95$, p<.0001, partial $\eta^2=.63$, $1-\beta=1$, while there was no main effect of Language Context, $F(1, 63)=.38$, p=.58, or interaction between the two factors, $F(1, 63)=2.6$, p=.11, suggesting that the self-bias effect was similar in the native- and foreign-language trials (4.95% vs. 6.67% of errors, respectively). Similarly, participants responded more accurately to the stimuli in the familiar than in the other condition, $F(1, 63)=33.31$, p<.0001, partial $\eta^2=.35$, $1-\beta=1$, and the Language Context effect was significant too, $F(1, 63)=7.33$, p<.05, partial $\eta^2=.1$, $1-\beta=.76$). However, there was no significant two-way interaction, $F(1, 63)=2.77$, p=.1.

The analysis of mismatching trials showed that these were responded to significantly more quickly in the native than in the foreign language (834 ms vs. 848 ms, respectively), t(63)=4.4, p<.0001. No significant differences were observed in the accuracy data (17.56% in the native- vs. 17.24% in the foreign-language context, respectively, t(63)=.24, p=.81).

Once the difference in the magnitude of the self-bias effect between native and foreign languages had been replicated, a regression analysis was carried out using participants' self-bias effects in English as the dependent variable, together with the list of 7 predictor variables derived from the questionnaires. We calculated the self-bias scores in English by subtracting the RTs in the self-condition from the RTs in the other-condition (M=113 ms, SD=73 ms; see Table 2). The predictor factors were the linguistic, affective and affective-linguistic variables described above: 1) *age of acquisition of English*, 2) *general English proficiency*, 3) *exposure to English*, 4) *empathy quotient,*5) *language-based emotional reactivity score*, 6) *self-perceived emotionality of English as a language*, and 7) *mean impact of emotional expressions in English* (see Materials and Procedure for more details; also see Appendix).

The regression model was not significant, F(7, 56)=1.08, p=.39, suggesting that the set of 7 independent variables did not collectively account for the self-bias in English (R^2=.12). As shown in Table 3, the relative contribution of each of the variables was markedly low, with very few exceptions. None of the affective, linguistic or affective-linguistic variables significantly accounted for the differences in the magnitudes of the self-bias effect in English (ts<1.5), except for a marginal effect of the *self-perceived emotionality of English as a language* (p=.057; see Table 3). Considering that the explanatory capacity of the general model was negligible, we interpret these results as a demonstration that there was no clear-cut relationship between the increased emotional reactivity towards English as a foreign language and the magnitude of the self-bias effect in that language.

-Insert Table 3 here-

Discussion

In Experiment 2 we wanted to look more closely into the notion of emotional distance promoted by foreign language contexts. We did so by exploring how the construct of language attitude shapes the self-bias effect in the foreign language. Specifically, we wondered to what extent, if any, the language attitude correlates with the self-perceived emotional distance for each of our participants. Is there a certain type of relation between the construct of emotional blunting in foreign languages and the perceived, self-reported, emotional distance towards said foreign languages?

It needs to be acknowledged that there was no significant relation between the emotional attachment we have towards our foreign languages and the size of the foreign language effects as measured by the magnitude of the self-bias in the foreign language (i.e., English). We found no significant relationship between the self-appointed foreign language attitude scores and the emotional reactivity elicited by said language. The results suggested that our participants' emotional attachment to each of the languages did not exert any influence on the magnitude of the effects.

General Discussion

The present study aimed to discern the fundamental core of the reduced emotional resonance driving the foreign language effects observed in the reduced self-bias. Across different designs, in our preceding study (Ivaz, Costa, & Duñabeitia, 2016) we obtained very robust yet ambiguous effects that were in line with the previous literature both on the self-bias and foreign language effects. However, it was left unclear which aspect of the non-native language (in our case, English) was at the root of these effects: the fact that English was not acquired from birth (i.e., its *non-nativeness*) or that it is a language and culture to which there is highly limited exposure and use (i.e., its *foreignness*).

To allow for teasing these aspects apart, we employed the same simple paradigm and tested our participants in their native, second, and foreign languages (Experiment 1). Overall, the participants exhibited significant self- and familiar-biases across all languages, replicating existing findings in the field (e.g., Sui, & Humphreys, 2015a, 2015b; Sui, He & Humphreys, 2012; Sui, Rotshtein, & Humphreys, 2013; Sui et al., 2014, 2015). Critically, the foreign language condition yielded a significantly diminished self-bias as compared to the native and non-native language conditions, both of which were virtually identical. The latter is congruous with the findings indicating that the emotional impact in a language is much greater when a language is acquired through immersion as opposed to through formal instruction (Caldwell-Harris, 2014; Dewaele, 2010). Our findings showed that the origin of the non-native language effects in the self-paradigm is indeed language foreignness and not non-nativeness alone, allowing these effects to retain their foreign language effects label. Negligible differences were found between the data in the native- and second-language conditions, indicating that the emotional responses in these languages are highly similar, if not equal. Thus, these findings suggest similar levels of emotional attachment to native and second languages, and as such, they go hand in hand with the reality of language use around us nowadays. It appears that the very experience of being surrounded by and immersed in a language, when combined with living amongst its community members, will invariably attach us emotionally to said language (Caldwell-Harris, 2014; Dewaele, 2010).

Sometimes, the emotional distance towards a language is associated with the physical distance from said language and its culture. For a number of reasons, this physical distance can disappear, either geographically (e.g., if we move to the country where our foreign language is contextually present), or mentally (e.g., if we are very invested in that culture, follow it closely and immerse ourselves in it). This brings up the question of whether this newfound physical/mental proximity could somehow implicate the emotional one by proxy. We argue that the answer to this question is unequivocal, affirmative, and relatively easily provable – simply by testing people who have

undergone this shift (e.g., migrants). This would be reflected in the language attitude of the very people who have experienced this during their lifetime, and would, therefore, be the one driving the foreign language effect. Our linguistic circumstances can change fairly easily according to several scenarios: by moving to a country where the spoken language is foreign to us, or falling in love with a native speaker of that language, or using that language on a daily basis in our workplace, or with our friends and loved ones, etc. Specifically, once we start having these different experiences that attach us to a language – albeit a foreign one – this language is no longer *foreign* to us. The emotional resonance is developed when the words of a language are either learned (in the case of native languages) or used (as in second languages) in emotional circumstances, because we process, remember, and make sense of our experiences based on their contexts (Caldwell-Harris, 2014). Much like what Marian & Kaushanskaya (2008) report with the bilingual immigration narratives they collected, we argue that emotional expression and emotional language seem to depend on personal experiences.

Experiment 2 investigated the role that perceived emotional proximity to a foreign language has in the foreign language effects. Once more we used the self-paradigm in the native- and foreign-language conditions. None of the variables obtained using the current set of questionnaires were able to account for the differential effect in the magnitude of the self-bias. While some of these factors were initially expected to majorly drive the effect, these results showed that it did not at all vary as a function of participants' reported emotionality with regard to the foreign language, their linguistic preferences or even their socio-linguistic context. Whether these variables were inadequately defined, ill-assessed or simply inaccurate, we do not know for certain. However, it is worth considering that perhaps there is something in the automaticity of the self-bias that transcends these variables which are relying on metacognitive judgements. Considering that none of the affective, linguistic or affective-linguistic variables included in the regression model of Experiment 2 seemed to correctly account for the foreign language effects in the self-paradigm, we

posit that one could potentially turn to the self-bias per se as a measurement of the emotional reactivity we have in a given language. The self-bias is an automatic, reliable and robust effect which is, as such, a quantitative, palpable measurement that could be used as an indicator of a person's emotional resonance in foreign language contexts.

Together, the results of these experiments demonstrate that the foreign nature of languages that are not our native tongues is at the core of the reduced emotional resonance tied to self-reflection in the self-paradigm. Moreover, these results suggest that the emotional distance towards our languages may not be a general construct, but rather a very elusive and difficult to gauge variable that is not necessarily reflected by the perceived emotional distance of each individual with regard to each of the known languages. In Experiment 2, we initially endeavored to uncover the affective factors which would best explain the differences in the magnitude of the self-bias across language contexts. However, we failed to account for this difference using the set of variables obtained in a series of questionnaires. In light of the results, we reconsidered our position and now tentatively recommend the use of the self-bias itself as a possible measurement of the emotional attachment towards the languages we know.

Finally, one should also keep in mind the specific characteristics of the self-paradigm that make it well-suited to explore emotional reactivity in the absence of confounding factors. A reason why the self-paradigm lends itself well to researching the foreign language effects is its extremely low linguistic content. As opposed to the most common materials for this type of research, i.e., emotional words and phrases, it is worth noting that the self-paradigm we used has 3 high-frequency common and short words (e.g., Sui, & Humphreys, 2015a, 2015b; Sui, He & Humphreys, 2012; Sui, Rotshtein, & Humphreys, 2013; Sui et al., 2014, 2015). Their low emotional content notwithstanding, these materials still allow for obtaining robust foreign language effects (Ivaz, Costa, & Duñabeitia, 2016). This allowed us to circumvent the issue of the cognitive cost associated

with foreign language processing, and the extent to which it figures in the foreign language effects observed in our results. That way, the foreign language effects we obtained were not due to our participants being less proficient in their foreign language than in their native and second languages. Consequently, we posit that the foreign language effects obtained are due to the emotional distance as a result of the linguistic context and not to the cognitive cost incurred by the processing of a non-native language, be it second or foreign.

This self-paradigm we employed could be utilized in several different ways in future research by adjusting the types of stimuli while still tapping into language-related (or more broadly framed, culture-related) questions. Namely, it especially lends itself to research that aims to converge culture effects and in-group effects. For instance, images that can be associated with different languages or cultures (i.e., flags, tourist attractions, cultural symbols, prominent figures, etc.) could be employed in this paradigm to test a variety of emotion-driven, automatic biases that represent different forms of linguistic or cultural relativity.

It is important to drive home the main point of the present two experiments: the foreign language effects, at least as reflected in the reduced self-bias effect, are undoubtedly present in the case of foreign and not second languages, and cannot be easily predicted in terms of a number of socio-linguistic and affective factors, including our perceived emotional distance towards our languages. We conclude that perhaps the self-bias itself, here used as an overt measurement of the foreign language effects, could possibly be used as a way of assessing our emotional resonance in different languages. Concerning the broader implications of this study, the differential effects pertaining to self-bias could have a bearing on the field of education and educational policies regarding curricula in foreign languages. The more we learn about the nature of the foreign language effects, the interplay of emotional resonance and the way we encode our concepts according to it, the better equipped the scientific community will be to weigh in on these educational policies (see Duñabeitia,

2017, for discussion). There is still much debate over whether teaching different school subjects in

foreign languages is the most beneficial choice for children, and more research into the foreign

language effects is warranted to help elucidate this issue.

References

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders, 34*(2), 163–175.

Bond, M. H., & Lai, T. M. (1986). Embarrassment and code-switching into a second language. *Journal of Social Psychology, 126,* 179-186.

Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology, 75*, 219-229.

Caldwell-Harris, C. L. (2008). Language research needs an "emotion revolution" AND distributed models of the lexicon. *Bilingualism: Language and Cognition, 11,* 169-171.

Caldwell-Harris, C. L. (2014). Emotionality differences between a native and foreign language: Theoretical implications. *Frontiers in Psychology: Language Sciences, 5,* 1055-1058.

Corey, J. D., Hayakawa, S., Foucart, A., Aparici, M., Botella, J., Costa, A., et al. (2017). Our moral choices are foreign to us. *Journal of Experimental Psychology: Learning, Memory and Cognition, 43*(1).

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.

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apestguia, J., Heafner, J., & Keysar, B. (2014). Your Morals Depend on your Language. *PLoS One*. e94842.

de Bruin, A., Carreiras, M., & Duñabeitia, J. A. (2017). The BEST dataset of language proficiency. *Frontiers in Psychology, 8*:522.

de Langhe, B., Puntoni, S., Fernandes, D., & van Osselaer, S. M. J. (2011). "The Anchor Contraction Effect in International Marketing Research." Journal of Marketing Research 48(2), 366-380.

Degner, J., Doycheva, C., & Wentura, D. (2011). It matters how much you talk: on the automaticity of affective connotations of first and second language words. *Bilingualism: Language and Cognition, 15*, 181-189.

Dewaele, J.-M. (2004a). Perceived language dominance and language preference for emotional speech: The implications for attrition research. In Schmid, M. S. & Köpke, B., Kejser, M. & Weilemar, L. (eds.) *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Amsterdam/Philadelphia: John Benjamins, pp. 81-104.

Dewaele, J.-M. (2004b). The emotional force of swearwords and taboo words in the speech of multilinguals. *Special issue of Journal of Multilingual and Multicultural Development, 25*(2-3), 204-222.

Dewaele, J.-M. (2009). Perception, attitude and motivation. In: Cook, V. and Li, Wei (eds.) *Language Teaching and Learning*. Contemporary Applied Linguistics 1. London, UK: Continuum, pp. 163-192.

Dewaele, J.-M. (2010). Multilingualism and affordances: Variation in self-perceived communicative competence and communicative anxiety in French L1, L2, L3 and L4. *International Journal of Applied Linguistics, 48,* 105-129.

Dewaele, J.-M. (2013). Multilingualism and Emotions. In C.A. Chapelle (ed.), *The Encyclopedia of Applied Linguistics*, Oxford, UK: Wiley-Blackwell.

Dewaele, J.-M., & Pavlenko, A. (2001-2003). *Web questionnaire Bilingualism and Emotions*.

University of London.

Dewaele, J.-M., & Pavlenko, A. (2002). Emotional vocabulary in interlanguage. *Language Learning, 52,* 263-322.

Diener, E., & Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal of Personality & Social Psychology, 68*, 653-663.

Duñabeitia, J. A. (2017). Emotional diglossia in multilingual classroom environments: A proposal. *Psychology and Cognitive Sciences – Open Journal, 3*(3), 74-78.

Duñabeitia, J. A., & Costa, A. (2015). Lying in a native and foreign language. *Psychonomic Bulletin & Review, 22*(4), 1124-1129.

Frings, C., & Wentura, D. (2014). Self-priorization processes in action and perception. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1737-1740.

García-Palacios, A., Costa, A., Castilla, D., del Río, E., Casaponsa, A., & Duñabeitia, J. A. (2018). The effect of foreign language in fear acquisition. *Scientific Reports, 8*:1157.

Geipel, J., Hadjichristidis, C., & Surian L. (2015a). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology, 59*, 8-17.

Geipel, J., Hadjichristidis, C., & Surian, L. (2015b). The Foreign Language Effect on Moral Judgment: The Role of Emotions and Norms. *PLoS ONE 10*(7): e0131529.

Gollan, T. H., Weissberger, G., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. Bilingualism: Language and Cognition, 15, 594-615.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67-81.

Harris, C. L. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development, 25,* 223-247.

Harris, C. L., Ayçiçegi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics, 24,* 561-578.

Harris, C. L., Gleason, J. B., & Ayçiçegi, A., (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In: Pavlenko, A. (Ed.), *Bilingual minds: Emotional experience, expression, and representation*. Multilingual Matters, Clevedon, United Kingdom.

Harzing, A. W. (2006). "Response Styles in Cross-national Survey Research A 26-country Study." International Journal of Cross Cultural Management 6(2), 243-266.

Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., & Keysar, B. (2017). Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. *Psychological science,* 0956797617720944.

Hernandez, A. E., & Meschyan, G. (2006). Executive function is necessary to enhance L2 lexical processing during picture naming: Evidence from fMRI. *Bilingualism: Language and Cognition*, *9*, 177-188.

Iacozza, S., Costa, A., & Duñabeitia J. A. (2017). What do your eyes reveal about your foreign language? Reading emotional sentences in a native and foreign language. *PLoS ONE 12*(10): e0186027.

Ivaz, L., Costa, A., & Duñabeitia, J. A. (2016). The emotional impact of being myself: Emotions and foreign language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *42*(3), 489-496.

Kahneman, D. (2003). A perspective on judgement and choice. *American Psychologist*, *58,* 697-720.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kernis, M. H., Grannemann, B. D., & Barclay, L. C. (1989). Stability and level of self-esteem as predictors of anger arousal and hostility. *Journal of Personality and Social Psychology, 56*, 1013-1022.

Keysar, B., Hayakawa, S., & An, S. G., (2012). The foreign language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science, 23*(6), 661-668*.*

Marian, V., & Kaushanskaya, M. (2008). Words, feelings, and bilingualism: Cross-linguistic differences in emotionality of autobiographical memories. *Mental Lexicon, 3*, 72-90.

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49*, 377-412.

Northoff, G., & Hayes, D. (2011). Is Our Self Nothing but Reward? *Biological Psychiatry, 63*, 1019-1025.

Pavlenko, A. (2005). "Ask each pupil about her methods of cleaning": Ideologies of language and gender in Americanization instruction. *International Journal of Bilingual Education and Bilingualism, 8*(4), 275-297.

Pavlenko, A. (2006). Bilingual selves. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (pp. 1-33). Clevedon, UK: Multilingual Matters.

Pavlenko, A. (2007). Autobiographic narratives as data in applied linguistics. *Applied Linguistics, 28*(2), 163-188.

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology, 47,* 405-428.

Puntoni, S., de Langhe, B., & van Osselaer, S. M. J. (2009). Bilingualism and the Emotional Intensity of Advertising Language. *Journal of Consumer Research, 35,* 1012-1025.

Shin, H. I., & Kim, J. (2017). Foreign Language Effect and Psychological Distance. *Journal of Psycholinguistic Research, 46*: 1339.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3-22.

Sui, J., & Humphreys, G. W. (2015a). Super-size me: Self biases increase to larger stimuli. *Psychonomic Bulletin & Review, 22*, 550-558.

Sui, J., & Humphreys, G. W. (2015b). More of me! Distinguishing self and reward bias using redundancy gains. *Attention, Perception & Psychophysics, 77*(8), 2549-2561.

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 1105-1117.

Sui, J., Liu, M., Mevorach, C., & Humphreys, G. W. (2015). The salient self: The left intraparietal sulcus responds to social as well as perceptual salience after self-association. *Cerebral Cortex, 25*, 1060-1068.

Sui, J., Rotshtein, P., & Humphreys, G. W. (2013). Coupling social attention to the self-forms a network for personal significance. *Proceedings of the National Academy of Sciences of the*

United States of America, 110, 7607-7612.

Sui, J., Sun, Y., Peng, K., & Humphreys, G. W. (2014). The automatic and the expected self: separating

   self- and familiarity biases effects by manipulating stimulus probability. *Attention,*

   Perception, & Psychophysics, 76(4), 1176-1184.

Tangney, J. P., & Fischer, K. W. (Ed.). (1995). *Self-conscious emotions: The psychology of shame,*

   guilt, pride, and embarrassment. New York: Guilford.

The Basque Government, Department of Education, Linguistic Policy and Culture. (2014). *Fifth*

   Sociolinguistic Map. Vitoria-Gasteiz, Central Publishing Service.

Table 1. Descriptive statistics of the linguistic and demographic factors of the different experimental groups tested in Experiments 1 and 2.

Experiments	Age (in years)	Basque AoA (in years)	English AoA (in years)	Basque proficiency (1-10)	English proficiency (1–10)	Basque interview (1–5)	English interview (1–5)	Basque vocabulary (out of 65)	English vocabulary (out of 65)	Exposure to Basque (% of time)	Exposure to English (% of time)
Experiment 1	24.04 (4.87)	3.54 (1.9)	5.93 (2.31)	7.13 (1.24)	6.32 (1.44)	3.68 (0.48)	3.32 (0.48)	45.93 (8.48)	42.75 (10.73)	19.29 (10.86)	12.14 (7.38)
Experiment 2	19.14 (1.76)	-	4.72 (1.61)	-	6.13 (1.61)	-	-	-	-	-	26.34 (19.61)

Note. Standard deviations are provided in parentheses. The age of acquisition (AoA) corresponds to the estimated age (in years) at which participants acquired a given language, according to their self-reports. The proficiency scores were calculated according to participants' self-ratings (on a 1–10 scale), while the interview scores correspond to the interviewer's assessment of participants' fluency in an interview (on a 1–5 scale). The vocabulary score corresponds to the number of correctly named pictures from a battery of 65 drawings. Exposure to a language corresponds to the self-reported percentage of time participants were exposed to said language in their daily lives.

Table 2. Mean Reaction Times (in ms) and Percentages of Errors in Each Condition Tested in

Experiments 1 and 2.
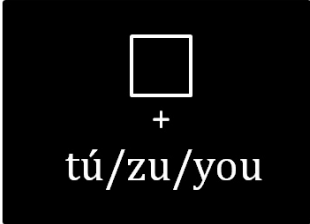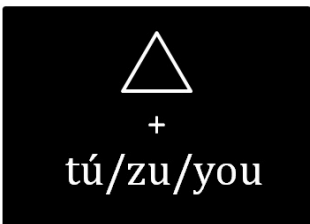
	Matching trials					Mismatching trials
Variables	**Self**	**Friend**	**Other**	**Self-bias**	**Familiar-bias**	
Native language Experiment 1						
RTs	621 (61)	696 (63)	731 (70)	110 ms	35 ms	748 (79)
% Errors	4.83 (4.33)	9.31 (9.9)	16.44 (12.11)	11.61%	7.13%	10.79 (6.44)
Second language Experiment 1						
RTs	626 (67)	692 (69)	737 (87)	111 ms	45 ms	761 (78)
% Errors	5.52 (5.37)	8.16 (6.21)	16.21 (12.93)	10.69%	8.05%	12.7 (7.08)
Foreign language Experiment 1						
RTs	645 (73)	708 (78)	722 (87)	77 ms	14 ms	758 (76)
% Errors	5.52 (4.82)	10.46 (9.03)	12.3 (10.08)	6.78%	1.84%	10.99 (8.16)
Native language Experiment 2						
RTs	657 (55)	759 (54)	791 (77)	134 ms	32 ms	820 (55)
% Errors	4.95 (6.34)	15.73 (11.86)	23.28 (14.47)	18.33%	7.53%	17.19 (10.08)
Foreign language Experiment 2						
RTs	674 (58)	760 (62)	786 (73)	113 ms	26 ms	829 (56)
% Errors	6.67 (6.29)	11.93 (9.39)	22.55 (14.62)	8.33%	5.23%	17.01 (9.89)

Note. Standard deviations are provided in parentheses. Self-bias effects are calculated by subtracting the RTs

and error rates in the self-conditions from those in the other conditions. Familiar-bias effects are calculated by

subtracting the RTs and error rates in the familiar conditions from those in the other conditions.

Table 3. Standard estimates (β coefficients), t and p values of the predictor variables included in the linear regression model tested in Experiment 2.

Predictor variables	Beta coefficient	t	p
English age of acquisition	-0.21	-1.48	0.14
English proficiency	-0.06	-0.41	0.68
Exposure to English	-0.06	-0.46	0.65
Empathy Quotient	-0.07	-0.54	0.59
Language-based emotional reactivity score	-0.12	-0.90	0.37
Self-perceived emotionality of English as a language	0.26	1.95	0.06
Mean impact of emotional expressions in English	-0.1	-0.74	0.46

Figure 1. Examples of stimuli from Experiments 1 and Experiment 2. The labels appeared either in Spanish, Basque or English, depending on the linguistic context of the experiment.

Figure 2. Schematic representation of the stimuli in Experiment 1 and 2.
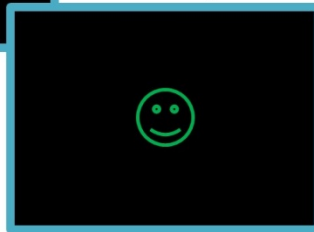
Fixation point (500 ms)

Critical stimuli
(1200-1500ms)

you

Feedback (500ms)

Appendix

The key questions from the questionnaire used to obtain the 3 different variables of interest that were used to classify the participants who took part in Experiment 2.

The 10 questions used to determine *the language-based emotional reactivity score* variable:

Please indicate which language (Spanish, English, or any other language) you would prefer to use in the following communicative scenarios:

1. If you were angry, in which language would you express your anger?

2. If you wanted to swear, in which language would you swear?

3. In which language would you express your deepest feelings?

4. When it comes to your internal monologue and consciousness, in which language do you form your inner thoughts?

5. Considering phrases like *Te amo* and *I love you*, in which language does this phrase have the most emotional impact for you?

6. If you kept a personal diary, in which language would you keep it?

7. If you had to speak about a bad memory or an embarrassing experience, in which language would you do it?

8. If your partner was a native English speaker and a proficient Spanish speaker as well, in which language would you talk to them?

9. In which language would you fight with your partner?

10. In which language would you praise or have an intimate conversation with your partner?

The 1 question used to obtain the *self-perceived emotionality of English as a language* variable:

1. Indicate how much you agree with the following statement: <u>English is emotional</u>. Please answer on a scale from "absolutely disagree" to "absolutely agree" (absolutely disagree, disagree, neither agree nor disagree, agree, absolutely agree).

The 9 questions used to determine the *impact of emotional expressions in English* variable:

1. How much of an impact does hearing or saying these words have on you in English: <u>reprimands and criticism</u>? Please answer on a scale from none to extreme (none, a little, quite, great, extreme).

2. How much of an impact does hearing or saying these words have on you in English: <u>insults and mockery</u>?

3. How much of an impact does hearing or saying these words have on you in English: <u>swearwords and expletives</u>?

4. How much of an impact does hearing or saying these words have on you in English: <u>taboo words</u>?

5. How much of an impact does hearing or saying these words have on you in English: <u>terms of endearment</u>?

6. How much of an impact does hearing or saying these words have on you in English: <u>greetings and congratulations</u>?

7. How much of an impact does hearing or saying these words have on you in English: <u>expressions of gratitude</u>?

8. How much of an impact does hearing or saying these words have on you in English: <u>compliments and flattery</u>?

9. How much of an impact does hearing or saying these words have on you in English: <u>advice</u>?