PROCESSING NEUTRAL, EMOTION AND TABOO WORDS: AN ERP STUDY
Master’s thesis, filed to obtain a Master’s degree in psychology
Specialization: theoretical and experimental psychology
by
Anneleen Mortier

Promotor: Prof. Dr. R. Hartsuiker
Supervisor: Elisah D’Hooge
Abstract

Bilinguals have reported that they have a stronger emotional experience in their native language (L1) than in their second language (L2). Behavioral studies are inconclusive about this question. However, studies that used skin conductance levels as a dependent variable are clearer about this effect: participants elicit a stronger skin conductance response on emotional and taboo words in L1 compared to L2 (e.g. Harris, Ayçiçegi and Gleason, 2003). Moreover, this skin conductance response was higher for taboo words than for negative words, indicating that taboo words are a special type of emotional words. This study tries to investigate the difference in emotional experience between L1 and L2 with a more directive and online measurement: ERP. The stimuli that are used are negative and neutral words (experiment 1) and taboo and neutral words (experiment 2). The dependent measures are the P300 and the long positive potential (LPP). Our hypothesis is that negative and taboo words elicit a P300. This P300 will be larger in L1 than in L2.

Key words: bilingualism, emotion, taboo, ERP, P300, LPP
María (L1: Spanish, L2: English):
‘I never swear in Spanish. I simply cannot. The words are too heavy and are truly a taboo for me.’
(Dewaele, 2004b:95)

The above quote illustrates how most of the bilinguals feel about their two (or more) languages: they experience greater emotion when using their native language (L1) than using their second language (L2) or other languages (LX). This phenomenon is experienced the strongest when bilinguals are using emotional or taboo words (e.g. Dewaele, 2004a, 2500g, 2008).

There are two possible views concerning this phenomenon. One side is the universal view: emotions are universal and exert an influence on language. People have emotions and are in an emotional state, this is independent of the language we speak. This emotional state can influence communication: the emotions can be seen extra linguistically (e.g. body language), but also lingual (= the things we say). In contrast with this, the cognitive and constructive approach believes that emotion is a side product of language. Because of our language, we give names to everything we see and feel, thus we also name emotions. Hence, we all talk about emotions. In this view, emotions do not influence language, but they are constructed by language. Thus, there can be a difference in how certain cultures feel about a certain emotions (e.g.: ‘pride’ is seen as positive in the US, but negative in India – “Emotions and Culture,” (2012)). Also, some terms of emotions only exists in one language (e.g. ‘schadenfreude’ in German, ‘saudade’ in Portugese – “Emotions and Culture,” (2012)).

So, these differences in experiencing a certain emotion show us that this might depend on the language we speak. Therefore, it is interesting to investigate this emotional value in different languages with bilinguals: will an emotional word have the same value in both languages (universal view) or will this value depend on the acquisition of the languages (constructivist view). The goal of this study tried to investigate whether there is an emotional difference between L1 and L2 in bilinguals. This was studied with negative and taboo words, as these are the two types of words that have the heaviest emotional connotation. We examined this effect with an objective, online and direct method – ERP signals.

In this study, we will discuss previous and more general research about multilinguals and emotions. Following, we will discuss the physiological studies with bilinguals and the emotional value of each language. Finally, we proceed to our own investigation.
Bilinguals and emotions

One of the strongest emotions that a human feels or experience is love. Thus, this emotion can differ in intensity when experienced or expressed in a different language. Dewaele (2008) administered an online questionnaire about experiencing love to a sample of 1459 multilinguals. They had to answer open and closed questions about the perceived emotional weight of ‘I love you’ in the different languages the participants spoke. He found that the majority of the respondents found this phrase to have the strongest emotional value in their native language (L1). However, this effect was dependent on certain other variables: age of acquisition, self-perceived language dominance and context of the acquisition of the second language (L2). Participants felt that the emotional weight is the strongest in their L1 and that no other language could match this feeling. Thus, multilinguals prefer to express one of the most powerful emotions in their L1.

Taboo words are also a special category of emotional words. Several studies have found their influence on different types of tasks such as memory tasks (e.g. Guillet, & Arndt, 2009; Jay, Caldwell-Harris & King, 2008), rapid serial representation task (e.g., Arnell & Mansfield, 2008), lexical decision (Zeelenberg, Bocanegra & Pecher, 2011) and language production (D’Hooge, De Baene, Severens & Hartsuiker, submitted; Severens, Janssens, Kuhn, Brass & Hartsuiker, 2011). Additionally, various studies have shown the effect of taboo words in different domains: enhanced attention (e.g. Eilola, Havelka & Sharma, 2007; Mathewson et.al., 2008), superior recall (e.g. Ayçiçegi & Harris, 2004) and heightened skin conductance responses (e.g. Harris, Ayçiçegi & Gleason, 2003). Taboo words are not used in any textbook; so instructed language learners attain them through interaction with native speakers. Additionally, these types of words are often the first words learned when acquiring L2. As a result, taboo words are learned in a similar situation in L2 as in L1. This means that bilinguals can have the same emotional experience to taboo words in L1 as in L2. However, this is not what Dewaele (2004 a, 2004 b) found. He administered other online questionnaires about the perceived emotional force of taboo words, swear words (Dewaele 2004a) and in what language bilinguals swear in (2004b). The first study showed that multilinguals found the emotional force of the taboo words and swear words were higher in their L1. This effect diminished for languages that are learned at a later time. In his second study, Dewaele found that multilinguals swear in their dominant language. This dominant language is determined by age of acquisition, frequency of use and the learned setting (naturalistic, (e.g.: parents), or instrumentally, (e.g.: classroom)). This dominant language is most of the time L1: participants found that L1 has more force to express themselves. Also, they felt more emotionally involved when using L1.
Others said that they could not swear in their native language because the emotional weight is too heavy or the words are too taboo. Some participants also mentioned that they could not express their emotional feelings in LX (= other language than the native language), because they cannot find the emotional load of that language. Additionally, multilinguals might swear in their other languages when they want to soften the force that those words are carrying or when they do not want to be associated with their expression.

Thus, it seems that taboo words evoke a heavier emotional experience in L1 than in L2. Additionally, participants indicated that they preferred to curse or swear in L1 because they understood the meaning better and because they had more emotional force. This illustrates that taboo words probably have a heavier emotional connotation than regular emotional words. Furthermore, studies found that taboo words had a better recall than negative words (Guillet & Arndt, 2009), and that taboo words caught more attention than negative words (Mathewson, Arnell & Mansfield, 2008). As a result, when a study investigates the emotionality in different languages (or just in a single language), it is not sufficient to only use emotional words, taboo words should be included as well.

The findings by Dewaele (2004a, 2004b and 2008) have been confirmed by clinical case studies: patients are more emotionally involved when they talk about their traumas in L1. Also, L1 can help to remember lost memories, in comparison to L2 (Buxbaum, 1949). If memories are retrieved in L2, they are colorless and lifeless. Additionally, when patients are talking about negative memories, they use their L2 more often, but when they recall warm and happy memories, they will use their L1 (Javier, 1996. For an overview: Schrauf, 2000). However, some bilinguals can experience detachment from their L1 when they underwent negative experiences (e.g. German with World War II). On the other hand, when people are more attracted to their ‘new’ language, they can experience this language as more emotional as their L1. This indicates that we cannot simply conclude that L1 is more emotional than L2: other factors can still influence this process.

According to our hypothesis, the findings of the clinical case studies and the studies of Dewaele (2004a, 2004b and 2008) can be summarized as: L1 is a personal language with emotional affiliation and is used to express feelings. In contrast with this, L2 is a more emotional distant language and is used when we do not want to be associated with that expression (Pavlenko, 2002). Bond and Lai (1986) confirmed this hypothesis. They let Chinese-English bilinguals discuss embarrassing topics in their L1 (Chinese) and L2 (English). The dependent measure was the total speaking time for each topic. Results showed that the bilinguals talked longer about embarrassing topics in L2 than in L1. Additionally, bilinguals switch to L2 to avoid anxiety-provoking stimuli (Javier & Marcos, 1989). This has also been
confirmed by Gonzalez-Reigosa, 1976: taboo words presented in L1 elicit more fear than presented in L2.

The clinical case studies and the studies of Dewaele (2004a, 2004b & 2008) already gave an indication on the perceived emotional differences between L1 and L2. However, the results of Dewaele were based upon self-report, which can be inaccurate. Also, clinical case studies were not representative for the entire population. Hence, the next studies are experimental investigations that explore the emotional difference between L1 and L2.

Studies with monolinguals have shown that we have a better and more detailed memory for negative words (e.g. Kensinger & Corkin, 2003) and for taboo words (e.g. Grosser & Walsch, 1966; Hadley & MacKay, 2006; Jay, Caldwell-Harris & King, 2008; MacKay et al., 2004, among others). Furthermore, as the above studies suggested, the recall can be influenced by the kind of language we use. The study of Anooshian and Hertel (1994) is one of the first who investigated this effect in bilinguals. They asked late Spanish-English bilinguals to rate emotional and neutral words presented in English or Spanish. The presentation of these words was blocked for each language. This was followed by a surprise recall task. Anooshian and Hertel found that their participants were better at recalling emotional words, but only when those words were presented in their L1. The authors suggested that when we learn a L2 at a later time, this wouldn’t be associated with emotional experiences and because of this, will be less emotionally laden. This supports the hypothesis of the current paper: words in a L1 will have a heavier emotional connotation than words in L2. Ayçiçegi and Harris (2004) tried to replicate this effect and included reprimands and taboo words in their stimulus set. They also mixed the presentation of the words in each language in each block. However, they found the opposite result with late Turkish-English (learned L2 at the age of 12 or 18 and migrated to the US at the average age of 22). Results showed that the participants were equally good at recalling negative English and Turkish words, but for some stimuli, this effect was stronger in the L2. Concerning taboo words, participants had better recall and recognition for these words in comparison with negative words. Ayçiçegi-Dinn and Caldwell-Harris (2009) did a follow up study: perhaps their findings could be related to the level of processing (shallow processing or deep processing). They had four conditions: word rating on emotional intensity (replication of previous study), counting letters (shallow), translation (deep) and word association (deep). Averaging over the four tasks, they found that recall was equally strong for L1 and L2, except for the reprimands (more recall in L2). For the taboo words, the same conclusions hold as their previous study (Ayçiçegi & Harris): participants had a better recall and recognition for them compared to negative and neutral words. Those two studies supports the suggestion that taboo words should be seen as a ‘special’ type of emotional words. Ferré, García, Fraga, Sánchez-Casas and Molero
(2010) investigated whether this effect was dependent of age of acquisition or the setting where bilinguals learned their second language (immersion context or instrumental context). Again, early Spanish-Catalan bilinguals (immersion context – experiment 2) and late Spanish-English bilinguals (instrumental context – experiment 3) had to rate emotional or neutral words. The language in which the words were presented was blocked (one block for each language). In both experiments, there was no difference in recall between the two languages. Also, when they compared both groups, there was no difference in the magnitude of the emotional effect in L2.

These four studies indicate that there is a great variability of results in the studies on memory recall in bilinguals. Anooshian and Hertel (1994) found better recall in L1, Ayçiçegi and Harris (2004) found better recall in L2 (for some stimuli), Ayçiçegi-Dinn and Caldwell-Harris (2009) and Ferré et.al. (2010) found no differences between the two languages. Furthermore, Ferré et.al. showed that language proficiency does not influence the emotional experience in L2. As the results of previous studies mentioned above, these suggest that bilinguals have a different emotional experience towards their L1 and L2, but the results from these memory studies are not consistent towards this effect. For taboo words, these studies show that participants have an advantage in recall for these words, but on average, there is no difference between L1 and L2. However, when words are encoded for storage, it could be that the words are being stripped from language, and that just the complex emotional term is stored. As a consequence, there are no differences between L1 and L2 in recall.

Another line of research is investigating attention towards emotional words. One of the most used paradigm to investigate selective attention is the Stroop task (Stroop, 1935). The emotional Stroop task is a modified version of the original Stroop task. In this task, participants are shown emotional and neutral words in different colors. Their task is to say the color of the words aloud, and ignore the written word. The general finding is that, when the target word is an emotional word, there is more interference when saying the color of the word (e.g. Pratto & John, 1991). This interference is caused by the emotional value the word has, and not by incongruency (as in the classical Stroop task). Sutton, Altarriba, Gianico and Basnight-Brown (2007) administered this emotional Stroop task by Spanish-English bilinguals. However, note that the bilinguals were more proficient in their second language than in their first. The results showed that there was an emotional Stroop effect: the participants were slower when they had to pronounce the color of an emotional word. Nevertheless, this effect was not moderated by language: there was no difference between L1 and L2. This could conclude that this null-effect could be due to the proficiency of the second language: the participants used their second language at that time more than their first, which could have altered their emotional experience to their L2. Eilola, Havelka and Sharma (2007) did an emotional Stroop task as well, but their
participants were late Finnish-English bilinguals. However, they were still very proficient in their L2. The authors used emotional and taboo words as distractors. Participants were slower on taboo words and negative words compared to neutral words. Likewise, they found no difference in effect between the first and the second language. Eilola et al. concluded that the emotional effect is dependent on the proficiency of the language: the more proficient you are, the more emotionally attached and the less discrepancy between L1 and L2 in the emotional Stroop task.

The authors suggested that proficiency does play a role in the emotional experience towards a certain language. However, this is in contrast with what Ferré et al. (2010) found. This again shows that there is no clear evidence towards what actually influences this effect when behavioral studies are used.

Another paradigm where the attention for emotional words is investigated is the rapid serial visual presentation (RSVP). This is a paradigm where stimuli are presented very shortly to participants. When distractor and target are presented within 500 ms from each other, an attentional blink can occur, resulting in worse performance on the target. Studies with monolinguals have already shown that when this distractor is an emotional or a taboo word, people have a slower reaction on the target (e.g. Zeelenberg, Bocanegra & Pecher, 2011). Moreover, Mathewson, Arnell and Mansfield (2008) found that taboo words led to more interference than emotional words (even negative words). Additionally, when the emotional word is a target, the attentional blink is reduced (Anderson, 2005). This suggests that emotional words capture the attention more than neutral words and that there are less attentional resources required to process emotional words. Colbeck and Bowers (2012) investigated this effect with bilinguals. They showed a stream of 18 words, with 1 critical distractor (English taboo or sexual word) and 1 color target to Chinese-English bilinguals and native English speakers. The participants had to identify the color-word. For the native English speakers, they found the classical finding: they were worse at identifying the color word when it was preceded by an emotional or a sexual distractor. For the bilinguals, this effect was reduced: they performed better on the task than the native speakers when there was an emotional or a sexual distractor. Even when they only took the ‘best scoring’ bilinguals in account, the effect still persists. These results are in contrast with the conclusions of Eilola et al. and Sutton et al.: it indicated that the first language is perceived more emotional than the second language, even when bilinguals are very proficient in their second language.

In short, these behavioral studies are more trustworthy than the self-reported studies from Dewaele (2004a, 2004b, 2008). These studies also suggest that there is a difference in emotional experience between L1 and L2, but this difference can be influenced by various
factors. Some of the above studies have proposed several variables that might influence this effect: e.g. age of acquisition, proficiency,… However, other studies dismissed these effects. The effects perceived in the results of the behavioral studies can also be influenced by other processes (e.g. attention, memory,…). Hence, it might be better to use a different dependent variable than reaction time. The brain-based perspective suggests that all subjective experiences have a measurable physiological correlate (Harris, 2004). Thus, physiological measurements can investigate the emotional difference between L1 and L2 in a purer and clearer way (e.g. this measurement can pick up effects that are not even conscious). Additionally, another advantage is that this can investigate the underlying mechanisms. Thus, the following studies investigated if bilinguals are more emotionally attached to their L1 using physiological correlates.

Bilinguals and physiological studies

The study of Harris et.al. (2003) is one of the first that investigated the emotional experience of bilinguals with skin conductance levels (SCL’s, formerly known as Galvanic skin response). They presented English and Turkish target words to Turkish-English bilinguals in the auditory or visual modality. These target words consisted of neutral words, positive and negative words, reprimands and taboo words. The reprimand were childhood reprimands such as ‘Shame on you!’. The language of the words was mixed. Results showed that the taboo words elicited the strongest autonomic response; these were even stronger when the taboo words were presented in their L1. Another finding was that when the stimuli were presented in the auditory modality, the skin conductance level was higher than when they were presented in the visual modality. In contrast with this, when the words were presented in L2, there was no difference between modalities. These results support the hypothesis of this study. Harris (2004) found that age of acquisition mediated this relationship between emotional experience to a language and the skin conductance level. She showed the same English target words as in Harris et.al. with their Spanish translations to Spanish-English bilinguals. They had two groups of participants: early bilinguals (average age of acquisition: 3.7 years old) and late bilinguals (average age of acquisition: 7.9 years old). The skin conductance level was measured as well. Results showed that late bilinguals had a stronger skin conductance response when they saw childhood reprimands in their first language (L1 – Spanish). In contrast with this, early bilinguals had no difference in skin conductance response if the childhood reprimands were shown in their L1 or L2. Harris explained this by age of acquisition: the early Spanish-English bilinguals had acquired their second language at an early age. Also, they moved to the US at an
early age (average: 3.1 years old), thus they are very familiar with this culture and emotional value of it. Harris and Ayçiçegi-Dinn (2009) also used skin conductance responses, but they let Turkish-English bilinguals read Turkish and English true of false statements. They found a stronger skin conductance response when the participants read false statements and this response was even stronger for L2 (English). These findings supported the ‘double stressor’ account: for most participants it was harder to lie in L2, because speaking a second language requires more effort than speaking a native language. Moreover, lying also requires more effort and cognitive resources than speaking the truth. Thus, it is not sure what causes these stronger responses: it could be the result of the additional cognitive resources (‘I want to do my best, because it is in English, and I don’t want to look like a fool’), or it could be the result of a more emotional experience to their second language.

The studies above suggest that indeed there is a difference between L1 and L2 in terms of emotional experience. These differences are also easily picked up by a physiological measurement. This supports the assumptions that the behavioral studies cannot pick up the underlying mechanisms that evoke this difference. This is also illustrated in the study of Eilola and Havelka (2010). They did a follow up on their previous emotional Stroop experiment (Eilola et.al., 2007). Native English speakers and Greek-English bilinguals had to react to the color of an English target word, and ignore the meaning of it. However, what was new to their experiment was that they measured skin conductance levels. The behavioral results showed no difference between monolinguals and bilinguals, but the SCL’s clearly showed a difference: it was significantly higher for the monolinguals in the condition of emotional and taboo words than for the Greek-English bilinguals. This study is very crucial because it suggests that at the behavioral level, there is no difference between the emotional perception of L1 and L2 (no difference in RT), but the physiological measurements are suggesting the opposite. Thus, it is very important to study the underlying mechanisms.

To summarize these studies, the physiological studies show more clarity: overall, emotional and taboo words in L1 elicit stronger skin conductance responses than in L2. Moreover, as Eilola et.al. (2010) showed, this is independent from behavioral results. However, skin conductance levels cannot give information why these levels are higher in L1 than in L2 (as in Harris and Ayçiçegi-Dinn (2009)). As the authors suggested, it could simply be explained by the emotional strength of L1. Nevertheless, SCL is not time-locked to the stimuli and is a more diffuse measurement, so more general and non-task related factors could influence this response (e.g. mood of the participant, external conditions,...). Moreover, SCL measures the reaction of the sympathetic nervous system. Thus, it would be better to study the origin of this reaction (the brain). A last remark is that the measure for SCL habituates when stimuli are
repeated (Codispoti, & De Cesarei, 2007). This habituation does not happen with ERP (LPP – Codispoti, & De Cesarei, 2007). Also, ERP is a continuous, direct and online measurement. It makes it possible to investigate every change between the stimulus presentation and the response. Another advantage of ERP is that this measurement is time-locked to the processing of the stimuli, so we know with certainty that an event has elicited that response in the ERP’s. Thus, we can be more conclusive about why these differences between L1 and L2 exist. As a conclusion, this is a better measurement to investigate these differences.

Present study

Emotional stimuli have the property to automatically capture our attention (see previous studies: e.g. Altarriba et.al., 2007; Zeelenberg, et.al., 2011). This attentional bias is also reflected in the ERP: when an emotional stimulus is presented, larger peaks within the 200 – 300 ms range appear (e.g. D’Hooge, et.al. (submitted); for an overview, see Hajcak, MacNamara & Olvet, 2010; Schupp et.al., 2006). Additionally, when presenting emotional stimuli, the P300 blends into the LPP. This is a broad superior-posterior positivity (Hajcak et.al., 2010). The difference between the P300 and the LPP is that the P300 is seen as a reflection of the increase in attention towards task-relevant stimuli. The LPP on the other hand is a reflection of the vigilance and processing of intrinsically motivating stimuli (Hajcak et.al., 2010).

Previous studies already showed that ERP is effective at measuring the emotional value of a certain language. Wu and Thierry (2012) used ERP in a translation-priming paradigm. Chinese-English bilinguals were presented English and Chinese emotional and neutral words. The first word served as a prime, the second one they had to judge whether it was related in meaning to the prime. The crucial part of their experiment was that for some words pairs, prime and target shared the same sounds when they were translated into Chinese, but they were unrelated in meaning and form in English (e.g. failure – poet : shi bai – shi ren). The ERP results showed that when prime and target are related in sounds in Chinese, the ERP’s in the negative valence condition was significantly different from the ERP’s in the positive and neutral conditions. Moreover, the negative valence failed to elicit the ERP amplitude. These results suggest that bilinguals use the L1 (Chinese) translation to judge whether the target was related to the prime, but this only happened with neutral and positive words. This indicates that negative valenced words have a ‘special status’ when compared to the positive valenced words: we are trying to protect ourselves against the negative connotation of these words by not
activating their meaning in our L1. This might indicate that this meaning is too emotional or too offensive in our L1.

Opitz and Degner (2012) investigated the different sense of emotionality in L1 and L2 with ERP. They presented positive, negative and neutral nouns to German-French and French-German bilinguals. The participants had to make a lexical decision on the target words. Their dependent measure was the early posterior negativity (EPN). What they found was no difference in amplitude for the EPN in L1 and L2, but it was delayed in L2. However, this experiment has some shortcomings, so that the results might be distorted. The authors matched their stimuli pairwise, but it has not been mentioned if this was also done for every condition. Previous studies have shown that the frequency can influence the amplitude: words with a low frequency can elicit a smaller P300 and high frequency words can elicit a smaller P1. For the EPN, low frequency neutral words elicit a larger EPN than high frequency neutral words. High frequent negative words elicit a larger EPN than low frequency negative words. Moreover, this difference is not observed for positive words (Scott, O’Donnell, Leuthold & Sereno, 2009). So, frequency can influence the amplitudes of the EEG. Thus, the difference between the amplitude for positive and negative words could be due to the frequency difference than to the emotionality between the two languages. A second remark is that they used stimuli from a database according to valence and arousal rated by monolinguals (German and French). When a study tries to investigate the emotionality of words in L1 and L2, several words will be more emotional in L1 than the translation equivalent in L2. They tried to solve this problem by using a balanced bilingual design. However, it would have been better if they have rated their stimuli by a similar sample of the population on valence and arousal to control for this effect. Another remark is that in their ERP, they only looked at the effect on the early time course. However, there can be some differences later in that time course (e.g. in the LPP). As a last remark, the authors used a lexical decision task. However, the retrieval of the word can influence the effect the word has when it was presented. Thus, the results of their study could be biased by factors they did not control. The present study tried to control for these confounds.

The goal of this study was to investigate with ERP whether there is a difference in emotional experience in L1 and L2. We used a delayed-naming task to investigate this matter, since this is the most neutral and less demanding task. Also, the emotion of the word is not needed in this task: it was activated automatically. Since we also wanted to investigate the difference between negative and taboo words, our critical stimuli were negative and neutral words (experiment 1) and taboo and neutral (experiment 2) words. The reason why we did not include any positive stimuli is because of the difference in processing between positive and negative words: e.g. subjects are faster on positive words compared to negative and neutral
words (Kissler & Koessler, 2011; Palazova, Mantwill, Sommer & Schacht, 2011). Also, we investigated the difference between taboo and emotional words in L1 and L2. Thus, negative words are more in line with the taboo words. The dependent measures were the P300 and the long positive potential (LPP). Our hypothesis was that negative and taboo words would elicit a P300. This P300 would be larger in L1 than in L2.

**Experiment 1**

Taboo words are a special type of emotional words. Several studies have confirmed this suggestion. Jay et al. (2008) have administered several memory tasks. They found that participants were better at recalling taboo words than negative words, even when they had to report those taboo words as last. Guillet & Arndt (2009) also found that participants had a better memory for taboo words than for negative words. However, they used a different task than Jay et al. (2008). In all their experiments, participants had to remember certain words (central information) in the presence of a distracter (peripheral information). This distracter could be either taboo or negative. In all of their experiments, they found that the central information was remembered better when the distracter was a taboo word, not when it was a negative word. In a study of Mathewson et al. (2008), they used a single-target rapid serial visual representation task (RSVP). Their results showed that the taboo words received more attention than negative words.

With ERP, we would like to investigate whether there is also a difference in brain activity between taboo and negative words and whether this is modulated by language. Jay et al. (2008) showed already that there is a physiological difference between taboo and negative words: taboo words elicited more frequently a skin conductance response than other words. However, ERP is a more sensitive and online method as skin conductance responses, so we chose to investigate this difference with ERP.

**Method**

**Participants.** 20 psychology students from Ghent University participated in this experiment (2 male and 18 female). They received a small compensation. Before they were accepted to this study, they had to fill out two questionnaires to make sure that all participants had a sufficient knowledge of English (L2). Another questionnaire was administered to determine their right-handedness. The first questionnaire they have to rate their L1 (= Dutch), L2 (= English) and L3 (= French) on a 7-point Likert scale, ranging from very bad to very good.
They all reported that Dutch was their native language (L1), and the language they used the most. Participants had a mean of 6.50 for L1, 5.45 for L2 and 3.70 for L3. (table 1). All participants received English (L2) education in secondary school and they lived in a Dutch dominant environment. All of them still come in contact on a regular basis with English by internet, tv, films, radio, etc. For their L2, most of the participants learned their this at home (n = 16) and some of them at school (n = 2). The other 2 participants learned English on holidays. The average age of acquisition for English was 5.47 years (ranging from 6 years old until 14 years old) and for French 3.82 years (ranging from 1 year old until 13 years old). This is no surprise, since French is an official national language and toddlers are already encouraged to learn French. Nevertheless, 17 participants reported English as their L2. The other 3 participants reported the same proficiency for English and French.

The second questionnaire was the LexTALE in Dutch and the LexTALE in English (Lemhöfer & Broersma, 2011). Only participants that had a score higher than 85% on Dutch and 60% on English were allowed to participate. Also, participants that scored higher on the English version than on the Dutch version were excluded. The average score on the Dutch version was 93.62% and for the English version 78.44%. To make sure we didn’t receive atypical brainwaves, all participants that had a neurological or psychiatric disorder were excluded from the study.

As last, participants had to fill out the Dutch handedness questionnaire (Van Strien, 1992; Van Strien, 2002). The questionnaire describes ten actions which participants had to judge with what hand they executed these actions (left hand, right hand, both hands). 16 participants were extreme right-handed (‘10’: maximum score for right handedness), 3 participants had a score of 9 for right-handedness and 1 participants had a score of 8, all indicating that they were right-handed.
<table>
<thead>
<tr>
<th>Language</th>
<th>Skill</th>
<th>Experiment 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Writing</td>
<td>6.20 (.53)</td>
</tr>
<tr>
<td>L1 (Dutch)</td>
<td>Reading</td>
<td>6.70 (.47)</td>
</tr>
<tr>
<td></td>
<td>Understanding</td>
<td>6.60 (.60)</td>
</tr>
<tr>
<td></td>
<td>Speaking</td>
<td>6.50 (.69)</td>
</tr>
<tr>
<td></td>
<td>General proficiency</td>
<td>6.50 (.61)</td>
</tr>
<tr>
<td>L2 (English)</td>
<td>Writing</td>
<td>4.89 (.88)</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>5.75 (.64)</td>
</tr>
<tr>
<td></td>
<td>Understanding</td>
<td>5.75 (.64)</td>
</tr>
<tr>
<td></td>
<td>Speaking</td>
<td>5.40 (.75)</td>
</tr>
<tr>
<td></td>
<td>General proficiency</td>
<td>5.16 (.51)</td>
</tr>
<tr>
<td>L3 (French)</td>
<td>General proficiency</td>
<td>3.70 (1.22)</td>
</tr>
</tbody>
</table>

Table 1: Self-Assessed Ratings (7-point Likert Scale) of L1, L2 and L3 Proficiency (Experiment 1). Standard deviations are included in parentheses.

**Stimuli.** Before we administered the experiment, we chose to rate our stimuli. Volunteers, who received a small contribution, performed this rating. The volunteers were similar to the subjects in the actual experiment in terms of education, age and L1 and L2 knowledge. First, a huge pool of potential emotional nouns was gathered, using the ANEW and previous studies (Ayçiçegi-Dinn & Caldwell-Harris, 2009; Belezza, Greenwald & Banaji 1986; Bertels, Kolinsky & Morais, 2009; Bertels, Kolinsky & Morais, 2010; Eilola, Havelka & Sharma, 2007; Ferré, Garcia, Fraga, Sanchez-Casus & Molero, 2010; Guillet & Arndt, 2009; Harris, Ayçiçegi & Gleason, 2003; Harris, 2004; Janschewitz, 2008; Pavlenko, 2002; Thomas, Johnstone & Gonsalvez, 2007 and Zeelenberg, Bocanegra & Pecher, 2011). English words were being translated to Dutch and vice versa. Then, we omitted all cognates and homophones from our stimulus set, because those words could enhance the bilingual lexical processing (e.g. Costa, Caramazza & Sebastian-Galles, 2000). Further, using the SUBTLEX-Dutch (Keuleers, Brysbaert & New, 2010) and SUBTLEX-English (Brysbaert & New, 2009), we searched for control neutral words, which are matched on frequency, word category (nouns) and number of letters (identical). Using a paired t-test, we found no significant difference between all categories of words in terms of frequency (p > .1). Eventually, only 50 words each category remained (in total: 200). Those stimuli were randomized for each participant in our rating study, so that every participant received a different order. Participants had to judge the stimuli on a 7-point likert scale on valence (1 = very negative, 7 = very positive) and arousal (1 = not arousing,
To make sure we had negative stimuli, we selected all Dutch negative words that had a rating lower than 2.5. For the neutral words, Dutch words were selected that had a rating higher than 4 and lower than 5. We only selected on the Dutch words, because else we would have restricted the effect: English and Dutch words would have the same judgments and the effect would be even. From our rated data pool, 26 stimuli per category were selected. Those are the critical stimuli that were being used in the delayed naming task. They are still matched in terms of frequency, number of letters and word category (p > .1) (see table 2). For the full list of critical items: see appendix A.

<table>
<thead>
<tr>
<th></th>
<th># Letters</th>
<th>p value letters</th>
<th>Word Freq</th>
<th>p –value freq</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negative</strong></td>
<td>7.27 (2.32)</td>
<td>7.00 (2.08)</td>
<td>&gt; .5</td>
<td>2.63 (.54)</td>
</tr>
<tr>
<td><strong>Control</strong></td>
<td>7.27 (2.32)</td>
<td>7.00 (2.08)</td>
<td>&gt; .5</td>
<td>2.72 (.55)</td>
</tr>
<tr>
<td><strong>p –value</strong></td>
<td>Identical</td>
<td>Identical</td>
<td>&gt; .25</td>
<td>&gt; .6</td>
</tr>
</tbody>
</table>

Table 2: Standard deviations are displayed in parentheses. Reported p-values: paired t-test between different conditions.

**Procedure.** Before the participants begun with the experiment, they were asked to fill out the informed consent. Next, they were being prepared for the ERP-experiment.

Participants sat in front of a computer screen, with a voicekey to respond. We emphasized on the fact that they should not blink during the trial (from fixation cross until after the response). At the beginning of the experiment, Dutch instructions were presented on the screen. The participants could not operate the computer on any condition.

The subjects started with one practice block of 20 trials (10 emotion words and 10 control words, different from the actual targets) in Dutch. Afterwards, three experimental blocks followed, also in Dutch. Each block contained the same words, but in a different randomized order, with the condition that no more then 3 words of the same category are allowed to be presented in a consecutive order. When the three experimental blocks were finished, another practice block started, this time with 20 English trials (10 emotion and 10 control words, translations from the first practice block). Again, three experimental blocks followed. The language blocks were contra-balanced over participants. Between the practice block and the experimental blocks and in-between the different blocks, there were short breaks where the experimenter could check and encourage the participant.

Every trial started with a fixation cross presented in the middle of the screen for 700 ms. Immediately the target word was presented for 1000 ms. Next, an exclamation mark was shown in the center of the screen, which was the cue for the participants to pronounce the word they
have just seen. This remained on the screen until the participants pronounced the words, or until the maximum response time was expired (= 5000 ms). The intertrial interval was 1500 ms, which is a sufficient amount of time to blink.

After the experiment, the participants had to make backwards and forwards translations on the critical stimuli. Also, they had to judge each word on emotionality on a 7-point likert scale (1 = negative, 4 = neutral, 7 = positive).

EEG recording and analyses. Four bipolar electrodes were placed below and above the eyes, to register vertical eye-movements (blinking). 2 other bipolar electrodes were placed next to eyes, to register horizontal eye-movements. The reference point was the 2 electrodes placed on the mastoid. The Biosemi system was used for the EEG. This was recorded continuously at a sampling rate of 512 $\Omega$ from 62 tin electrodes place according to the international 10-20 setting (Jasper, 1958). Electrical resistances were kept below 3K$\Omega$ at all electrodes. We referred offline to the average of the left and right mastoids; there was no online reference. Only an offline bandpass filter was used of .01 – 30 Hz.

EEG data were analyzed using EEPROBE (ANAT inc.). ERP’s were time-locked to the onset of the word, with epochs being generated in a time window running from -100 ms to 1000 ms. Before averaging, the epochs that contained errors (e.g. speech errors, verbal hesitations, EEG artifacts,…) were removed. The eye-movements were being corrected based on the substraction of the vertical EOG propagation factors, based on PCA transformed EOG components (Nowagk & Pfeifer, 1996).

Experiment 2

In the second experiment, we wanted to investigate if there is any difference in the emotional experience when participants see a taboo word in either Dutch or English. Also, with those two experiments, we can compare the differences in experience between emotional and taboo words. Again, we set up a rating study to get proper stimuli. We also used the same procedure and method as in experiment 1.

Method

Participants. 20 psychology students from Ghent University participated in this experiment (2 male and 18 female). They received a small compensation. Before they were accepted to this study, they had to fill out two questionnaires to make sure that all participants had a sufficient knowledge of English (L2). Another questionnaire was administered to
determine their right-handedness. The first questionnaire they have to rate their L1 (= Dutch), L2 (= English) and L3 (= French) on a 7-point Likert scale, ranging from very bad to very good. They all reported that Dutch was their native language (L1), and the language they used the most. All participants received English (L2) education in secondary school and they lived in a Dutch dominant environment. All of them still come in contact on a regular basis with English by internet, tv, films, radio, etc.

The second questionnaire was the LexTALE in Dutch and the LexTALE in English (Lemhöfer & Broersma, 2011). Only participants that had a score higher than 85% on Dutch and 60% on English were allowed to participate. Also, participants that scored higher on the English version than on the Dutch version were excluded. To make sure we didn’t receive atypical brainwaves, all participants that had a neurological or psychiatric disorder were excluded from the study.

As last, participants had to fill out the Dutch handedness questionnaire (Van Strien, 1992; Van Strien, 2002). The questionnaire describes ten actions which participants had to judge with what hand they executed these actions (left hand, right hand, both hands).

Stimuli. Stimuli were gathered from previous studies (Ayçiçegi-Dinn & Caldwell-Harris, 2009; Bertels, Kolinsky & Morais, 2009; Eilola, Havelka & Sharma, 2007; Guillet & Arndt, 2009; Harris, Ayçiçegi & Gleason, 2003; Harris, 2004; Janschewitz, 2008; MacKay et.al., 2004 and Zeelenberg, Bocanegra & Pecher, 2011) and from websites that indicated some words as taboo and inappropriate. English words were being translated to Dutch and vice versa. Next, using the SUBTLEX-Dutch (Keuleers, Brysbaert & New, 2010) and SUBTLEX-English (Brysbaert & New, 2009), we searched for control neutral words, which are matched on frequency, word category (nouns) and number of letters (identical). Again, our stimuli were rated by volunteers who received a small contribution. Participants had to judge the stimuli on a 7-point likert scale for personal use (1 = never use this word, 7 = use this word all the time), offensiveness (1 = not offended, 7 = very offensive), familiarity (1 = never encountered, 7 = encounter this all the time), tabooess (1 = not taboo, 7 = very taboo), valence (1 = very negative, 7 = very positive), and arousal (1 = not arousing, 7 = very arousing).

Procedure. Before the participants begun with the experiment, they were asked to fill out the informed consent. Next, they were being prepared for the ERP-experiment.

Participants sat in front of a computer screen, with a voicekey to respond. We emphasized on the fact that they should not blink during the trial (from fixation cross until after the response). At the beginning of the experiment, Dutch instructions were presented on the screen. The participants could not operate the computer on any condition.
The participants started with one practice block of 20 trials (10 taboo words and 10 control words, different from the actual targets) in Dutch. Afterwards, three experimental blocks followed, also in Dutch. Each block contained the same words, but in a different order. When the three experimental blocks were finished, another practice block started, this time with 20 English trials (10 taboo and 10 control words, translations from the first practice block). Again, three experimental blocks followed. The language blocks were contra-balanced over participants. Between the practice block and the experimental blocks, there was a short break where the experimenter could check and encourage the participant.

Every trial started with a fixation cross presented in the middle of the screen for 700 ms. Immediately the target word was presented for 1000 ms. Next, an exclamation mark was shown in the centre of the screen, which was the cue for the participants to pronounce the word they have just seen. This remained on the screen until the participants pronounced the words, or until the maximum response time was expired (= 5000 ms). The inter-trial interval was 1500 ms, which is a sufficient amount of time to blink.

After the experiment, the participants had to make backwards and forwards translations on the critical stimuli. Also, they had to judge each word on emotionality on a 7-point likert scale (1 = very negative, 4 = neutral, 7 = very positive). This procedure is exactly the same as in experiment 1.

**EEG recording and analyses.** Four bipolar electrodes were placed below and above the eyes, to register vertical eye-movements (blinking). 2 other bipolar electrodes were placed next to eyes, to register horizontal eye-movements. The reference point was the 2 electrodes placed on the mastoid. We used the Biosemi system for our EEG. This was recorded continuously at a sampling rate of 512 Ω from 62 tin electrodes place according to the international 10-20 setting (Jasper, 1958). Electrical resistances were kept below 3KΩ at all electrodes. We referred offline to the average of the left and right mastoids; there was no online reference. Only an offline bandpass filter was used of .01 – 30 Hz.

EEG data were analysed using EEPROBE (ANAT inc.). ERP’s were time-locked to the onset of the word, with epochs being generated in a time window running from -100 ms to 1000 ms. Before averaging, the epochs that contained errors (e.g. speech errors, verbal hesitations, EEG artifacts,…) were removed. The eye-movements were being corrected based on the substraction of the vertical EOG propagation factors, based on PCA transformed EOG components (Nowagk & Pfeifer, 1996).
References


Appendix A
Critical items in experiment 1

<table>
<thead>
<tr>
<th>English</th>
<th>Control</th>
<th>Dutch</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abuse</td>
<td>Pulse</td>
<td>Misbruik</td>
<td>Hartsig</td>
</tr>
<tr>
<td>Addict</td>
<td>Switch</td>
<td>Verslaafde</td>
<td>Schakelaar</td>
</tr>
<tr>
<td>Avalanche</td>
<td>Broadcast</td>
<td>Lawine</td>
<td>Omroep</td>
</tr>
<tr>
<td>Bullet</td>
<td>Border</td>
<td>Kogel</td>
<td>Grens</td>
</tr>
<tr>
<td>Crime</td>
<td>Entry</td>
<td>Misdaad</td>
<td>Toegang</td>
</tr>
<tr>
<td>Damage</td>
<td>Weight</td>
<td>Schade</td>
<td>Belang</td>
</tr>
<tr>
<td>Danger</td>
<td>Amount</td>
<td>Gevaar</td>
<td>Aantal</td>
</tr>
<tr>
<td>Despair</td>
<td>Descent</td>
<td>Wanhoop</td>
<td>Afkomst</td>
</tr>
<tr>
<td>Disease</td>
<td>Current</td>
<td>Ziekte</td>
<td>Stroom</td>
</tr>
<tr>
<td>Disgust</td>
<td>Bracket</td>
<td>Afkeer</td>
<td>Beugel</td>
</tr>
<tr>
<td>Dungeon</td>
<td>Lobster</td>
<td>Kerker</td>
<td>Kreeft</td>
</tr>
<tr>
<td>Fear</td>
<td>Meat</td>
<td>Angst</td>
<td>Vlees</td>
</tr>
<tr>
<td>Grief</td>
<td>Speed</td>
<td>Verdriet</td>
<td>Snelheid</td>
</tr>
<tr>
<td>Humiliation</td>
<td>Expectation</td>
<td>Vernedering</td>
<td>Verwachting</td>
</tr>
<tr>
<td>Injury</td>
<td>Player</td>
<td>Letsel</td>
<td>Speler</td>
</tr>
<tr>
<td>Manslaughter</td>
<td>Refrigerator</td>
<td>Doodslag</td>
<td>Koelkast</td>
</tr>
<tr>
<td>Massacre</td>
<td>Corridor</td>
<td>Bloedbad</td>
<td>Doorgang</td>
</tr>
<tr>
<td>Misery</td>
<td>Mirror</td>
<td>Ellende</td>
<td>Spiegel</td>
</tr>
<tr>
<td>Mutilation</td>
<td>Replacement</td>
<td>Verminking</td>
<td>Vervanging</td>
</tr>
<tr>
<td>Paralysis</td>
<td>Publisher</td>
<td>Verlamming</td>
<td>Uitgeverij</td>
</tr>
<tr>
<td>Plague</td>
<td>Button</td>
<td>Pest</td>
<td>Knop</td>
</tr>
<tr>
<td>Poison</td>
<td>Cattle</td>
<td>Gif</td>
<td>Vee</td>
</tr>
<tr>
<td>Rage</td>
<td>Soil</td>
<td>Woede</td>
<td>Bodem</td>
</tr>
<tr>
<td>Resignation</td>
<td>Inheritance</td>
<td>Ontslag</td>
<td>Erfenis</td>
</tr>
<tr>
<td>Suffocation</td>
<td>Hibernation</td>
<td>Suffocation</td>
<td>Winterslaap</td>
</tr>
<tr>
<td>Violence</td>
<td>Evidence</td>
<td>Geweld</td>
<td>Bewijs</td>
</tr>
</tbody>
</table>