BRIEF REPORT

Blinded by Taboo Words in L1 but Not L2

Katie L. Colbeck and Jeffrey S. Bowers
University of Bristol

The present study compares the emotionality of English taboo words in native English speakers and native Chinese speakers who learned English as a second language. Neutral and taboo/sexual words were included in a Rapid Serial Visual Presentation (RSVP) task as to-be-ignored distracters in a short- and long-lag condition. Compared with neutral distracters, taboo/sexual distracters impaired the performance in the short-lag condition only. Of critical note, however, is that the performance of Chinese speakers was less impaired by taboo/sexual distracters. This supports the view that a first language is more emotional than a second language, even when words are processed quickly and automatically.

Keywords: emotion, second-language speakers, attentional blink

Bilingual speakers often report that their first language (L1) feels more emotional than their second language (L2). This has been documented in a wide variety of contexts, including writings of bilingual authors, testimonials of psychotherapy patients, interviews, and surveys (e.g., Dewaele, 2004a, 2004b). Several experimental findings also support this conclusion. For example, Harris, Ayçiçeği, and Gleason (2003) measured the *electrodermal-activity* (EDA) response to taboo words (EDA measures skin conductance, i.e., sweat) in Turkish-English bilinguals who acquired English later in life (after age 12). The authors observed greater EDA responses to L1 compared to L2 words, suggesting that L1 words had a greater emotion impact.

However, L1 words are not always more emotional. For example, in contrast with the above mentioned EDA results, Spanish-English bilinguals who learned L2 English early in life showed no difference in EDA to L1 and L2 taboo words (Harris, 2004; Harris et al., 2006). Another demonstration that words in L1 and L2 can have the same emotional impact was reported in an adapted version of the Stroop task, the so-called emotional Stroop task. In this task, participants name the ink color of emotional and neutral words, and the standard finding is that color naming is delayed for emotional words (e.g., MacKay et al., 2004). Critical for present purposes, is that highly fluent Spanish-English bilinguals who learned both languages before the age of seven showed the same amount of interference in both languages (Sutton, Altarriba, Gianico, & Basnight-Brown, 2007).

One obvious interpretation of the above findings is that the emotional impact is related to the age at which words are acquired, with strong emotions associated with either L1 or L2 words that

are acquired early. However this cannot be the complete story, as there are also demonstrations that late-acquired words can evoke strong emotions as well. For instance, Ayçiçeği-Dinn and Caldwell-Harris (2009) carried out an episodic memory experiment in which emotional and neutral words were presented in Turkish and English. Across a number of study conditions, episodic memory was better for emotion words in Turkish-English bilinguals who learnt their L2 English in middle school. That is, the emotional attributes of words had a similar impact on memory in the two languages (also see Ferré, Garcia, Fraga, Sanchez-Casas, & Molero, 2010). Some additional insight into the conditions in which emotions are connected to words in L1 and L2 comes from studies that have assessed bi- and multilinguals' responses to web questionnaires. Although most self-reports indicated that L1 is more emotional, the studies also reveal that the age at which L2 words are learned, the context in which L2 words are learned (e.g., naturalistic context vs. classroom setting), or the time away from L1 (when L1 can undergo language attrition), can all impact on the relative emotional weighting of L1 and L2 (cf., Dewaele, 2004a, 2004b, 2005, 2010). For an excellent review, see Pavlenko (2008).

One finding that seems to stand out from the above was reported by Eilola, Havelka, and Sharma (2007). They reported equally large emotional Stroop effects in unbalanced, bilingual Finnish-English participants who learned English late, and who for the most part acquired their L2 in a classroom setting. Specifically, participants were slower to name the color of the ink for negative, positive, and taboo words in L1 and L2 (compared to neutral words). Because the participants were all quite fluent in English, Eilola et al. argue that proficiency alone is the critical factor in this task, with the age and context of learning L2 words irrelevant.

How can these findings be reconciled with previous work suggesting that L2 words learned under these conditions are less emotional than L1 words? Eilola et al. claim that task differences may be critical. Specifically, they note that the emotional Stroop task provides an automatic measure of word processing and it discourages participants from focusing on the meaning of words

(they simply have to name the ink color). By contrast, studies that have found differences in emotion to L1 and L2 words have tended to ask participants to focus on the meaning of single words (e.g., Harris et al., 2003), or considered the emotionality of words in more complex contexts (e.g., web questionnaire studies in which participants reflect on their emotional reaction to words). Accordingly, different tasks may be tapping different aspects of emotionality in bilingual language processing. Specifically, Eilola et al. suggest that the automatic processing of emotions may only be a function of language proficiency, with age and context of learning L2 words irrelevant to performance on an emotional Stroop task. By contrast, more elaborative and strategic processing of emotions may be a function of all these factors.

The current experiment attempts to assess the claim that automatic processing of emotion words is simply the product of proficiency, with little impact of the age or context in which an L2 was learned. Specifically, we used an RSVP task and procedure that has been used in the past to study the automatic activation of emotions by words in L1. However, we adapted this procedure to the bilingual context.

In the RSVP task, words are rapidly and sequentially presented in the same spatial location. Typically, items are presented for a duration that allows participants to identify a target relatively accurately. A classic finding showed that when searching for two targets presented within 500 ms of each other, accuracy in reporting the second target (T2) was reduced, resulting in a so-called attentional blink (AB; Broadbent & Broadbent, 1987). Anderson (2005) showed that attentional blink was reduced when T2 was an emotional word. This suggests that fewer attentional resources are required for the processing and identification of emotional words. Indeed, Anderson and Phelps (2001) found that patients who had the left amygdala (which is involved in the processing of emotional stimuli; LaBar & Phelps, 1998) removed do not show this reduced AB, reinforcing the emotional saliency explanation. In addition, Mathewson, Arnell, and Mansfield (2008) also found that the AB was *larger* when T1 was a taboo word, and they took these findings to support the claim that taboo words are more effective in capturing attention.

The present experiment is based on studies that have shown that AB can occur as a result of non-searched-for distracters in the RSVP stream (Arnell, Killman & Fijavz, 2007). The authors presented negative, positive, sexual/taboo, neutral, threat, and anxiety to-be-ignored distracters to participants at long and short lags before the target word. The taboo/sexual words resulted in an AB, suggesting that the taboo/sexual words capture the attention of native speakers even when attention is not directed to them. However, only taboo/sexual words presented a short lag before the target resulted in an AB, therefore attention capture only lasted for 500 ms.

In the present investigation we replicate the basic procedure of Arnell et al. (2007), in which a target color word was presented in an RSVP sequence, which was preceded by either a taboo/sexual or neutral word. The key difference is that we compared the performance of L1 and L2 speakers of English. The question is whether the performance of the L2 speakers of English (native Chinese participants) is affected by single English taboo/sexual words in comparison to L1 English participants. Specifically, if English taboo/sexual words have a similar emotional impact on native English and Chinese speakers, then both groups should

show a similar AB effect in the short lag condition. By contrast, if English taboo/sexual words have less emotional impact on the Chinese speakers, then the AB effects should be reduced in the Chinese participants.

Method

Participants

Twenty-four Chinese participants (8 male and 16 female), with ages ranging from 18 to 34 years ($M = 21.7$ years), and 20 native English speakers (8 male and 12 female), with ages ranging from 19 to 34 years ($M = 20.4$ years) voluntarily participated in the study. Participants had normal or corrected-to-normal vision. All participants individually completed a single 30-min session. Chinese participants were living in England to study for degrees taught in English, and were all proficient English speakers. More details about their competence will be described below.

Design

Three independent variables were manipulated. The first variable, critical distracter type, was a within-participants factor and had two levels: taboo/sexual and neutral. Equal numbers of taboo/sexual and neutral critical distracters were presented to each participant. There were 26 words in each condition (See Appendix A). A critical distracter was present on every trial. The second variable, lag length, was also a within-participants factor. Lag length represented how many items before the target the critical distracter was presented. It had two levels: close (a lag of three words) and far (a lag of eight words). Equal numbers of taboo/sexual and neutral critical distracter words were presented at each lag length. The third variable, nationality, was a between-participants factor and had two levels: English and Chinese. Both Chinese and English participants took part in all 208 trials, and trials were presented in a unique, fully randomized order to each participant.

Apparatus and Stimuli

The experiment was run using the DMDX software (Forster & Forster, 2003) installed on a Dell XPS laptop computer. All words were capitalized and presented in black 18-point Courier New font using RSVP, where stimuli were sequentially presented in the same spatial location. Each word was presented in the center of a uniform white screen for 100 ms, with no interstimulus interval between words.

Fifty-two critical distracter words (26 taboo/sexual, 26 neutral) were taken from Arnell et al., (2007). Both taboo and sexual words were used in the taboo/sexual condition. Sixty nonwords from four to seven letters in length were created for the noncritical distracters. This ensured that the neutral and taboo/sexual critical-distracter words were equally novel compared to the noncritical distracters. Therefore, only the ability of the critical distracters to capture attention, and not their comparative novelty, was compared.

Eighteen items were presented in each RSVP stream: 16 noncritical-distracter nonwords, 1 critical-distracter word (taboo/sexual or neutral), and 1 color target word (*black, blue, brown, green, orange, pink, purple, silver, white,* or *yellow*). The target

words were presented in stream position eight, 11, 13 or 16. The critical distracter was presented in stream position five or eight, three or eight items before the color target word. For the participants' responses, the color words were affixed to the keys of a USB-connected number pad.

The words used in the lexical-decision task were all low frequency (Kučera–Francis written frequency of between five and ten per million) and were between four and seven letters in length. The nonwords were based on words, but included a letter substitution or exchange. Nonwords were all pronounceable.

Procedure

Participants were asked to identify the color word presented on each trial by pressing the appropriately labeled key on the USB number pad, guessing if necessary, and to ignore all other words in the stream. They were shown the 10 color words and informed that the target would always be from this set and that only these responses could be given. Participants were told that emotional words may be presented on some of the trials, but that they should ignore these and focus on the color words. Five practice trials, which contained no critical distracters, preceded the experimental trials. Each trial began with the presentation of a black fixation cross in the center of the screen for 500 ms, followed by a 500 ms blank interval before the start of the RSVP stream. Each item in the RSVP sequence was presented for 100 ms. Immediately after each stream, participants were prompted by a sentence on the screen to press the key matching the identity of the target. Accuracy was stressed, and responses were not speeded. 2000 ms after the button press, the fixation cross for the next trial appeared. After the five practice trials, participants indicated that they had no questions and were ready to continue to the experimental trials by pressing the space bar. There was a self-timed break after half of the experimental trials had been completed.

Immediately after completion of the RSVP task, participants were given a tick list of all 52 critical distracter words and indicated how confident they felt defining the words (*completely, somewhat,* or *not at all*). This was used as a subjective measure of English proficiency. Participants also completed a 44-item lexical-decision task in which they categorized letter strings as English words or not. They were informed that although accuracy was important they should try to respond as quickly as possible. Each word remained on the screen for 2000 ms, or until the participant made a response. The first four trials were treated as practice trials and data from these were not analyzed. The lexical-decision task was used as an objective measure of participants' knowledge of English words.

Data Analysis

In order to be included in the experiment, participants were required to indicate on the tick list that they felt completely confident in defining at least 80% (42/52) of the critical distracter words. Five of the participants (four Chinese and one English) were excluded from the analyses as a result.

Results

First, a two-way repeated-measures ANOVA that included distracter type (taboo vs. neutral) and lag length (short vs. long) was carried out on the RSVP results of the English participants. More correct responses were recorded for the neutral (73.4%) than the taboo (65.4%) condition, $F(1, 19) = 30.56, p < .001$, and more correct responses were recorded in the far lag (72.3%) than the close lag (66.5%) condition, $F(1, 19) = 18.15, p < .001$. There was a significant interaction between word type and lag length, $F(1, 19) = 20.76, p < .001$, indicating that word type has different effects on the percentage of correct responses depending on the lag length. A Tukey post hoc test for pairwise comparisons between means showed that there was a significant difference between the taboo-close condition and all other conditions (p values $< .01$), and no other differences approached significance.

Next, a two-way repeated-measures ANOVA that included distracter type (taboo vs. neutral) and lag length (short vs. long) was conducted on the RSVP results of the Chinese participants. More correct responses were recorded in the neutral (73.3%) than in the taboo condition (69.8%), $F(1, 23) = 5.47, p = .028$. However there was no significant difference between the far lag (72.7%) and the close lag (70.4%) conditions, $F(1, 23) = 2.07, p = .164$. There was no significant interaction between word type and lag length, $F(1, 23) = 2.46, p = .130$.

To compare the RSVP results of the English and Chinese participants, a mixed ANOVA with two repeated-measures variables, word type (taboo vs. neutral) and lag length (close vs. far), and one between-groups variable, nationality (English vs. Chinese), was conducted. There was no main effect of nationality, with Chinese (74.3%) and English participants (72.1%) performing similarly overall $F(1, 42) = .469, p = .497$. There was a main effect of word type, with more correct responses recorded in the neutral (76.2%) than the taboo condition, (70.2%), $F(1, 42) = 29.79, p < .001$, and a main effect of lag length, with more correct responses in the long- (75.3%) compared to short-lag condition (71.1%), $F(1, 42) = 14.29, p < .001$. There was no significant interaction between lag length and nationality, $F(1, 42) = 2.71, p = .107$, reflecting the fact that English and Chinese participants were similarly affected by lag length. However, there was a significant interaction between word type and lag length, $F(1, 42) = 18.27, p < .001$. A Tukey post hoc showed that there was a significant difference between the taboo-close condition and all other conditions, $p < .01$, and no other significant differences between conditions.

The critical question was how the results of the Chinese participants compared to those of the English participants. There was a significant interaction between word type and nationality, $F(1, 42) = 4.54, p = .039$, reflecting the fact that English speakers did more poorly on the taboo/sexual word condition. There was also a significant interaction between word type, lag length and nationality, $F(1, 42) = 4.26, p = .045$, reflecting the fact that English speakers performed more poorly on the taboo/sexual words in the close-lag condition. Clearly, taboo/sexual words did have a greater impact on the English L1 speakers (see Figure 1).

One possible interpretation of these results is that Chinese participants did not know English sufficiently well to show as strong an attentional blink. In an attempt to address this concern we now consider the results from the lexical-decision task. Overall, Chinese participants had more incorrect answers (16.9%) than English participants (5.9%), $t(25.23) = 2.81, p = .009$, and English participants had a faster reaction time (RT) ($M = 766$ ms) than Chinese participants ($M = 983$ ms), $t(38) = 4.77, p < .001$. These
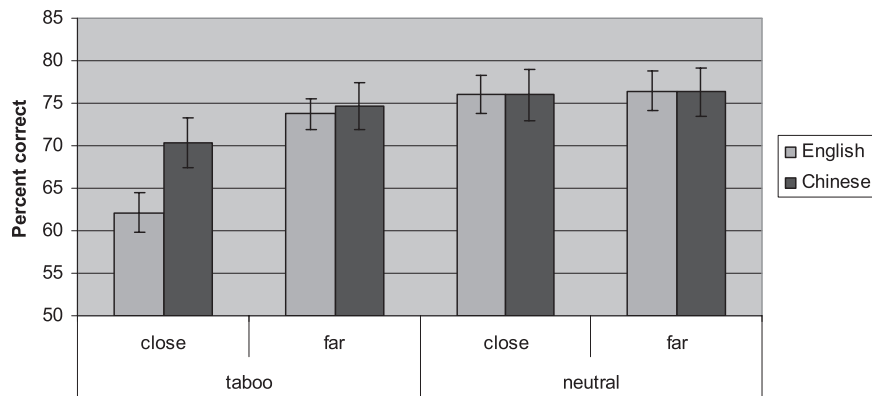
Figure 1. Percentage of correct color-target-identification responses as a function of critical distracter-emotional category and length of lag between the critical distracter and the color target for English and Chinese participants. Error bars indicate standard error.

relatively large differences in performance might appear to add some credence to this interpretation of cross-linguistic differences. We therefore reanalyzed the results of the Chinese participants who performed in the top 50th percentile in lexical-decision task (calculated by selecting the participants who had the lowest errors coupled with the fastest RT). These participants had an average of 6.1% incorrect and mean RT of 891 ms. Clearly their performance was still below native English speakers, but their level of English proficiency was quite high. Nevertheless the same pattern of results was obtained for this group, as shown in Figure 2.

In addition, we analyzed the RSVP performance as a function of the tick-list task in which participants rated their knowledge of the target words. English participants were completely confident in defining an average of 99.5% of the critical to-be-ignored distracter words. The Chinese participants were completely confident in defining an average 95.2% of these words; somewhat confident in defining an average of 2.8% of the words; and not at all confident in defining an average of 2% of the words. When only the performance of participants who felt completely confident in defining all the taboo words (11 Chinese and 17 English) was analyzed, the same overall pattern was maintained. In this case, the

difference in error rate between the neutral and taboo distracters in the near condition was 6.7% for the Chinese participants and 13.0% for the English participants (the corresponding difference for all participants was 5.6% and 13.9%, respectively).

Discussion

Consistent with the Arnell et al. (2007) findings, native English speakers were poorer at identifying a target word in an RSVP stream preceded by a taboo/sexual word than by a neutral word. That is, taboo/sexual words generated an attentional blink (AB). The critical new finding is that the AB to these taboo/sexual words was reduced in bilingual Chinese-English speakers who acquired their L2 English later in life. This was the case, even though all participants were fluent in English (they were each taking a degree course at the University of Bristol), and when 95% of the participants reported that they were completely confident that they could define the taboo/sexual words. Indeed, the same pattern of results was obtained when we restricted our analyses to those Chinese participants who were 100% confident in defining all the words, or when we restricted our analysis to the Chinese participants who
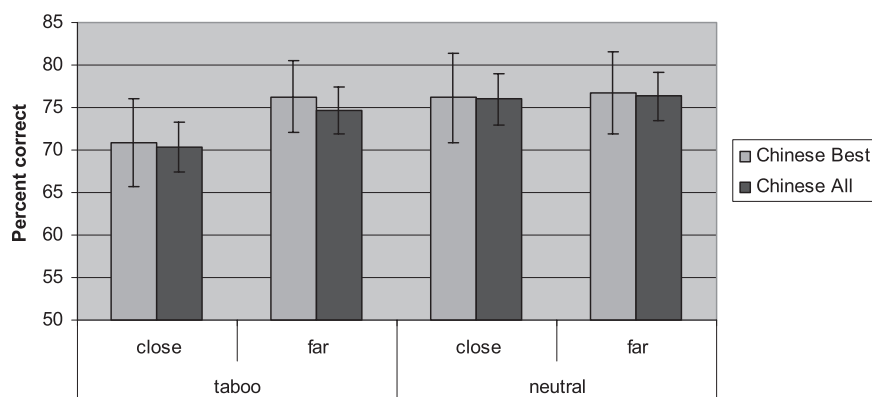

Figure 2. Percentage of correct color-target-identification responses as a function of critical distracter-emotional category and length of lag between the critical distracter and the color target for the best-performing Chinese participants and all Chinese participants. Error bars indicate standard error.

were in the top 50th percentile in an unrelated lexical-decision task.

The current results contrast with those of Eilola et al. (2007) who found that English taboo/sexual words are equally capable of capturing attention in English monolinguals and Finnish-English bilinguals who acquired their L2 English after the age of seven. They used their findings to argue that the automatic retrieval of emotions from L2 words is largely a function of proficiency, with the age and context of learning L2 largely irrelevant. On their view, age and context of learning an L2 can have an effect, but only in tasks that tap into later (more conscious) stages of word processing. That is, the authors are suggesting that the automatic and nonautomatic processing of L2 words is differently sensitive to emotional associations. Our finding that the attentional blink is reduced in response to L2 words challenges this conclusion.

It is difficult to know why the emotional Stoop and the attentional blink tasks produced different results with L2 taboo/sexual words. One possibility is that the Finnish-English participants studied by Eilola et al. were more proficient in English than our Chinese-English participants. Still, it is worth noting again that our study was carried out in an English context (students receiving a degree at the University of Bristol) whereas the Eilola et al. study was carried out in a Finnish context (Helsinki) with over 50% of their participants reporting that they had spend less than 6 months of their life in an English-speaking community. Nevertheless, we cannot rule out the possibility that the difference in outcomes was due to the fact that our participants were less proficient in English. Even so, our Chinese-English participants were proficient in English, and they clearly showed reduced emotional response to L2 words in an automatic word processing task.

In our view, the current findings challenge the conclusion that automatic and higher-level processing of L2 words are differentially sensitive to emotion. Indeed, our emotional-blink data are consistent with a range of studies that have found that the emotionality of L2 words is a function of a wide range of variables, including proficiency, the age at which the language was acquired, the context in which L1 and L2 were acquired, and the frequency with which the words are currently used (cf. Pavlenko, 2008). At present, the most parsimonious conclusion is that this holds for both the automatic and strategic processing of L2 words.

References

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdale impair enhanced perception of emotionally salient events. *Nature, 411*, 305–309. doi:10.1038/35077083

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General, 134*, 258–281. doi:10.1037/0096-3445.134.2.258

Arnell, K. M., Killman, K. V., & Fijavz, D. (2007). Blinded by emotion: Target misses follow attention capture by arousing distractors in RSVP. *Emotion, 7*, 465–477. doi:10.1037/1528-3542.7.3.465

Ayçiçe˘gi-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism: Language and Cognition, 12*, 291–303. doi:10.1017/S1366728909990125

Broadbent, D. E., & Broadbent, M. H. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics, 42*, 105–113. doi:10.3758/BF03210498

Dewaele, J.-M. (2004a). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development, 25*(2-3), 204–222. doi:10.1080/01434630408666529

Dewaele, J.-M. (2004b). Perceived language dominance and language preference for emotional speech: The implications for attrition research. In M. S. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First language attrition: Interdisciplinary perspectives on methodological issues* (pp. 81–104). Amsterdam, The Netherlands: John Benjamin.

Dewaele, J.-M. (2005). Investigating the psychological and emotional dimensions in instructed language learning: Obstacles and possibilities. *Modern Language Journal, 89*, 367–380. doi:10.1111/j.1540-4781.2005.00311.x

Dewaele, J.-M. 2010. *Emotions in Multiple Languages.* Basingstoke: Palgrave Macmillan. doi:10.1057/9780230289505

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion, 21*, 1064–1076. doi:10.1080/02699930601054109

Forster, K. I., & Forster, J. C. (2003). A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124.

Ferré, P., Garcia, T. Fraga, I., Sanchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion, 24*, 760–785. doi:10.1080/02699930902985779

Harris, C. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development, 25*, 223–247. doi:10.1080/01434630408666530

Harris, C. L., Ayçiçe˘gi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first than in a second language. *Applied Psycholinguistics, 24*, 561–579. doi:10.1017/S0142716403000286

Harris, C. L., Gleason, J. B., & Ayçiçe˘gi, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (pp. 257–383). Clevedon, England, UK: Multilingual Matters.

LaBar, K. S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science, 9*, 490–493. doi:10.1111/1467-9280.00090

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition, 32*, 474–488. doi:10.3758/BF03195840

Mathewson, K. J., Arnell, K. M., & Mansfield, C. A. (2008). Capturing and holding attention: The impact of emotional words in rapid serial visual presentation. *Memory & Cognition, 36*, 182–200. doi:10.3758/MC.36.1.182

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition, 11*, 147–164. doi:10.1017/S1366728908003283

Sutton, T. M., Altarriba, J., Gianico, J. L., & Basnight-Brown, D. M. (2007). The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion, 21*(5), 1077–1090. doi:10.1080/02699930601054133

(Appendix follows)

Appendix A

Taboo and Neutral Distracters

Neutral	Taboo
aisle	anus
binder	ass
butter	bastard
cable	bitch
card	boobs
chat	clitoris
chew	cock
computer	dick
cotton	dildo
desk	erotic
fish	fuck
gel	incest
glove	lesbians
guzzle	naked
haggle	naughty
jacket	nipples
justify	orgasm
loop	orgy
planet	penis
ruffled	piss
send	pussy
staple	sexual
vote	shit
walk	slut
wire	tits
zipper	whore