This is an Accepted Manuscript of an article published by Taylor & Francis in COGNITION AND EMOTION on April 9th 2025, available at: https://doi.org/10.1080/02699931.2025.2488986.

Early Identification of Taboo Words Reveals a Prominent Role of Semantic Information in Visual Word Recognition

Simone Sulpizio[1,2] and Michele Scaltritti[3]

[1]Dipartimento di Psicologia, Università degli Studi di Milano-Bicocca,

[2]Milan Center for Neuroscience (NeuroMI), University of Milano-Bicocca

[3]Dipartimento di Psicologia e Scienze Cognitive, Università degli Studi di Trento

Word Count: 5441

Short Title: TABOONESS BOOSTS WORD IDENTIFICATION

Author note

Simone Sulpizio: https://orcid.org/0000-0002-4051-9494

Michele Scaltritti: https://orcid.org/0000-0002-3328-1748

None of the authors has any conflict of interest to disclose.

Correspondence concerning the article should be addressed to: Simone Sulpizio, Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo, 1 – 20126 Milano, Italy, email: simone.sulpizio@unimib.it; Michele Scaltritti, Department of Psychology & Cognitive Science, University of Trento, Corso Bettini, 84 – 38068 Rovereto (TN), Italy, email: michele.scaltritti@unitn.it

Abstract

This research used the progressive demasking paradigm to investigate whether perceptual word identification is facilitated by semantic information. Experiment 1 revealed faster identification for taboo than neutral words. Experiment 2 revealed faster identification for taboo than emotionally-comparable non-taboo words, whereas the difference with respect to neutral words was possibly mitigated by list-wise factors related to list composition. Moreover, the facilitation for taboo words was impervious to habituation. The taboo connotation advantage seemingly originates from the attentional capture triggered by tabooness, a socio-culturally determined semantic feature that, under appropriate contextual conditions, modulates perceptual word identification. Our results suggest that a) semantic processing is a pervasive component of any task involving word processing, and b) when semantic information does not hinder the main task, it may influence even the earliest stages of word perceptual identification.

Keywords: taboo words; attention; attentional capture; progressive demasking; awareness; semantics

Extracting information from visual stimuli is fundamental to many of our interactions with the environment. This process depends on attentional dynamics, which modulate our subjective awareness of the surrounding stimuli (e.g., Lamme, 2003; Sergent et al., 2013) by selectively focusing our attentional resources on specific sources of information as a function of the ongoing goals. Certain stimulus properties are particularly salient and may thus benefit from a favored interaction with attentional resources. This is the case, e.g., of emotional stimuli (e.g., emotional faces, or threatening stimuli like a snake), which rapidly capture our attention (e.g., Hodsoll et al., 2011; Ohman & Mineka, 2001), thereby quickly reaching subjective awareness (e.g., Gayet et al., 2016; Yang et al., 2007), possibly due to their relevance for survival.

The relation between stimulus significance and attention becomes more ambiguous when considering symbolic stimuli. A paradigmatic example is language. From a phylogenetic perspective, words have no inherently relevant meaning, when considered as perceptual objects. For example, the word *snake* is not always associated with an immediate danger and can appear in several unthreatening contexts. Relatedly (and unlike non-symbolic emotional stimuli), emotional words have shown inconsistent effects on participants' behavioral performance. Specifically, emotional connotation has been associated to: a) both facilitation (e.g., Kousta et al., 2009) and interference in word recognition (e.g., Estes & Adelman, 2008), b) either enhanced (Anderson, 2005) or undistinguishable (e.g., Mathewson et al., 2008) attentional capture compared to neutral connotation, and c) earlier (Sklar et al., 2012) or equal (e.g., Rabagliati et al., 2018) availability to subjective awareness compared to neutral stimuli. These inconsistencies are even more noteworthy when considering the similarities across the studies reporting contrasting results. For example, both Kousta et al. (2009) and Estes and Adelman (2008) employed a visual lexical decision task, with positive and negative English words as stimuli. Thus, the effects triggered by emotional words appear

less consistent and more elusive than those observed with pictorial stimuli (e.g., Hinojosa et al., 2009; Kensinger & Schacter, 2006).

A different scenario emerges when considering a specific category of symbolic stimuli, that is taboo words. These are socially inappropriate terms including insults, slurs, sexual and scatological referents, as well as terms referring to physical or psychological deviations. Taboo words tend to be emotionally connotated, being typically characterized by high arousal (and often low valence; Sulpizio et al., 2024). Furthermore, their public use is restricted or even sanctioned, and it is precisely this prohibition that distinguishes taboo from other types of words (e.g., Jay, 2009). Research has shown that taboo words are processed more slowly than non-taboo ones in tasks involving word processing (e.g., Stroop paradigm: Bertels & Kolinsky, 2016; MacKay et al., 2004; lexical decision: Carretié et al., 2008; Madan et al., 2017; Scaltritti et al., 2021; Sulpizio et al., 2019; 2022; reading: Sulpizio et al., 2024). Critically, these interference effects have been linked to attentional capture: Because of its salience, the taboo connotation would operate as a prepotent semantic feature that captures participants' attention, at the expense of the ongoing task (e.g., Scaltritti et al. 2021; Sulpizio et al., 2022; for alternative interpretations, see, e.g., Dhooge & Hartsuiker, 2011; Williams & Evans, 1980). Clearly, this interpretation hinges upon the notions of a) an early semantic access during visual word recognition (e.g., Balota & Yap, 2016; Harm & Seidenberg, 2004; Pulvermuller et al., 2009; Sulpizio et al., 2022; Whiting et al., 2017), b) the relevance of attentional factors in visual word recognition (e.g., Besner et al., 2016), and c) the perspective of a flexible lexical processor (Balota & Yap, 2016), in which attentional-control systems adaptively weighs the emphasis on different information codes (i.e., orthographic, phonological, semantic) depending on situational demands.

Taboo-driven attentional capture has gained significant insights from studies exploiting taboo words to investigate the dynamics of temporal attention. For instance, the

attentional capture exerted by taboo connotation has been observed in studies examining the attentional blink phenomenon – a temporary deficit in reporting stimuli presented in rapid succession. In this paradigm, when a taboo word is the first of two consecutive to-be-reported targets, participants are more likely to miss the second target (i.e., tabooness induces a larger attentional blink), compared to when the first word is emotional (but non-taboo) or neutral. Instead, when a taboo word appears as the second target, participants are more likely to detect it (i.e., tabooness reduces the attentional blink; e.g., Anderson, 2005; Mathewson et al., 2008). Both phenomena highlight the attentional capture elicited by taboo words, suggesting that these stimuli demand a greater allocation of attentional resources, likely due to their high arousal and relevance (cf. Scherer, 2001; Schimmack & Derryberry, 2005). This heightened attention may, in turn, facilitate their access to awareness (Anderson, 2005; Schimmack, 2005; Lang et al., 1993).

A direct consequence of attentional capture and prioritized access to awareness would be an enhanced performance when the task goals focus on perceptual identification, due to an attention-driven boost in perceptual processing (e.g., Carrasco et al., 2004; Moray et al., 1976; Pinzmetal et al., 1998). Based on these premises, perceptual identification should be easier for taboo words than for non-taboo words. To test this hypothesis, the present study employed the progressive demasking paradigm (Grainger & Segui, 1990) across two experiments. In this paradigm, a mask (i.e., a series of #) and a word are alternated on the screen over a series of presentation cycles. All cycles have the same duration but, as the trial unfolds, the duration of the word progressively increases by a constant interval at the expense of the mask duration. Participants are instructed to press a button as soon as the word is identified. Crucially, the progressive demasking task primarily relies on visuo-perceptual identification processes (Grainger & Segui, 1990; Schreuder and Baayen, 1997), as it focuses on the exact recognition of the to-be-identified words. The identification (with an overt

report) of a unique lexical entry makes the task different from the lexical decision paradigm, in which responses are driven also by the familiarity with the stimulus (e.g., Balota & Chumbley, 1984), the summed activation of different lexical candidates within the lexicon (e.g., Carreiras et al., 1997; Grainger & Jacobs, 1996), and, more generally, any information that becomes promiscuously available to the decision mechanism beyond the activation of the target entry in the orthographic lexicon (Mahon & Caramazza, 2008). Also, compared to reading aloud, the progressive demasking task does not require an overt vocal response, which may emphasize the dimension of social inappropriateness and related interference effects (e.g., Sulpizio et al., 2024).

In the present study, we presented taboo words and compared them with non-taboo ones. We hypothesized that, in the context of an experimental task hinging upon recognition under conditions of limited visibility, taboo words would be recognized earlier than non-taboo words due to an attention-driven boost in perceptual processing. Such a finding would highlight the interactions between semantics and attention during visual word identification.

As a second step, we explored whether the expected facilitation induced by taboo connotation is subject to habituation. Habituation to taboo effects has been reported across various paradigms, with taboo interference decreasing over the course of the experiment (e.g., Stroop: MacKay et al., 2004; lexical decision: Sulpizio et al., 2022; reading aloud: Sulpizio et al., 2024). However, such effects have not been observed in paradigms that more directly assess attentional processes (as, e.g., the attentional blink; Arnell et al., 2007; Mathewson et al., 2008), where the attentional capture elicited by taboo words appears to persist over relatively long periods. Based on this evidence, we hypothesized that any potential facilitation for taboo words would remain stable across trials without habituation. This prediction, driven by previous evidence from attentional tasks, is also grounded in the notion that habituation functions as a filter for distracting yet irrelevant information (e.g., Cowan,

1999), preventing limited cognitive resources from being allocated to goal-irrelevant information. Clearly, while this filtering mechanism may contribute to the attenuation of interference effect in tasks where taboo connotation hinders performance (e.g., lexical decision or Stroop tasks; e.g., Sulpizio et al., 2022; Scaltritti et al., 2022), it may not apply when the taboo connotation enhance processing, as expected in the present study.

We tested our hypotheses in two experiments using the progressive demasking procedure. In Experiment 1, we used taboo and neutral words, with the two categories of stimuli being comparable for the main psycholinguistic properties except for affective properties (i.e., valence and arousal). In Experiment 2, we included an additional set of control words alongside the stimuli from Experiment 1, plus filler words. Control words were comparable to taboo words also for affective properties, but did not carry the social restrictions associated with taboo words. Filler words were comparable to neutral words for affective properties, and were included to keep the same proportion of affectively and non-affectively connotated words, exactly as in Experiment 1. By including control words, we aimed to isolate the specific contribution of taboo connotation in attentional-driven perceptual facilitation, over and above affective features.

Experiment 1

Methods

Participants

Forty students of the University of Milano-Bicocca took part in the experiment (31 females, mean age = 22.60; SD = 1.76). They were all Italian native speakers, and reported normal or corrected-to-normal vision and no history of learning disabilities. Participants received course credits for their participation. The study was evaluated by the local commission for minimal-risk studies of the Psychology Department at the University of Milano-Bicocca (protocol nr.: RM-2020-237).

The sample size was decided on the basis of recent recommendations in the field (Brysbaert, 2019; Brysbaert & Stevens, 2018), indicating that adequate power within repeated measure RT experiments can be attained with (at least) 1,600 observations per condition (our experiments had 2,800 observations per condition). Additionally, sample size is consistent with our previous in-lab experiments on taboo words (Scaltritti et al., 2021; Sulpizio et al., 2024) and with previous experiments using progressive demasking (e.g., Duñabeitia et al., 2008; Ferrand et al., 2011). Data are fully available at osf.io/3tfzc

Stimuli

Seventy taboo words and 70 non-taboo words were selected. Taboo words (e.g., cazzo, *dick*) were selected from ITABOO (Sulpizio et al., 2020), a database containing Italian taboo words together with ratings of several semantic dimensions. Neutral (non-taboo) words were selected from the Italian adaptation of the Affective Norms for English Words database (Montefinese et al., 2014). Taboo and neutral words were selected in order to be comparable across the main psycholinguistic variables known to affect word processing. Taboo and neutral words differed on valence and arousal, with taboo words being both more negative and more arousing than neutral ones (see Table 1).

Table 1

Psycholinguistic variables of the stimuli used Experiment 1.

Variables	Taboo	Neutral	*p* value
Frequency (log)	8.90	9.63	.14
Imageability	6.28	6.63	.13
Concreteness	6.47	6.61	.53
Valence	3.96	6.07	<.001
Arousal	5.18	4.48	<.001
N. of Letters	7.23	7.19	.87
OLD20	2.16	2.02	.24

Note. OLD = orthographic Levenshtein distance (Yarkoni et al., 2008) and Frequency values (log-transformed) were taken from the SUBTLEX-IT database (Crepaldi et al., 2013). Concreteness, imageability, valence, and arousal scores were taken from ITABOO (Sulpizio et al., 2020) and the Italian adaptation (Montefinese et al., 2014) of the Affective Norms for English Words database (Bradley & Lang, 1999) for taboo and nontaboo words, respectively.

Apparatus and Procedure

The experiment was controlled via the E-Prime 3 software (version 3.0.3.82) running on a computer desktop. Participants sat in front of the computer screen at a distance of about 60 cm, and were instructed to press the spacebar as soon as they could recognize a word.

All stimuli were presented in a single block, in a fully randomized order. Stimuli were presented in lower-case on the centre of the screen, in 18-point in Courier-new font, in black against a white background. Each trial cycle was composed of a word immediately followed by a mask of hash marks of the same length. During the trial, the word-mask sequence was sequentially repeated and, at each repetition, the duration of both the word and the mask was jointly manipulated by slightly increasing the word duration, and slightly decreasing the mask duration. During the trial, the total duration of the word-mask pair was held constant at 333 ms. At each successive iteration, the duration of the word increased by 13 ms (one refresh cycle, 75 Hz) whereas the duration of the mask decreased by the same amount. Thus, within each trial, at the first presentation the word and the mask were displayed for 13 ms and 320 ms respectively, at the second presentation they were displayed for 26 ms and 307 ms, and so on. The sequence was repeated until the participant response or until when the word was presented for all the 333 ms interval. Response latencies were measured from the beginning of the first word-mask pair presentation until the participant's response. After their response, participants were also asked, with no time limit, to type in the word they have just seen. In case of mistyping, they could delete and type the word again. Participants could initiate the next trial by pressing the spacebar. The whole experiment lasted ~15 minutes.

Statistical analyses

Reaction times (RTs) were analyzed via linear mixed effects models and response accuracy via generalized linear mixed effects models, using the *lme4* library (version 4_1.1-21; Bates et al., 2015) in R (version 4.2.2, R Core Team, 2022). Figures were made using *ggplot2* package (version 3.4.1, Wickham, 2016). Word category (taboo vs neutral) was included as fixed factor, and participants and items as random intercepts. Fixed effects were assessed via likelihood ratio tests, by comparing the model in which the fixed effect under examination was present vs. the model in which it was absent. RTs analyses were conducted on correct responses only, whereas accuracy analyses were conducted on the whole dataset.

Results

Two participants were discarded, one because of excessive errors (i.e., mistyped words, > 80% of responses) suggesting very low task compliance, and one because of excessively slow response latencies (more than 2.5 *SD* from the sample mean).

The pattern of RTs is represented in Figure 1. The analysis of RTs showed that word category approached significance (χ^2 [1] = 3.81, p = .05), with taboo words being recognized faster than neutral ones (b = -64.83, SE = 33.12, t = -1.96). To further test the reliability of this finding, we refitted the model after having removed as outliers those observations displaying a residual error above 2.5 *SD* (e.g., Baayen et al., 2008). The refitted model showed a fully significant effect of word category (χ^2 [1] = 5.30, p = .02; b = -69.93, SE = 30.20, t = -2.31)[1]. We also estimated the Bayes Factor (BF) for the effect of word category. Specifically, we subtracted the Bayesian Information Criterion (BIC) of the model featuring the fixed effect of word category from the one taken from the null model (only random intercepts), thus obtaining the delta BIC. The BF was then computed following the formula

[1] Note that the same model run on log-transformed RTs (and without outlier removal) showed again a significant effect of word category (χ^2 [1] = 4.94, p = .02, b = -0.036, SE = 0.016, t = -2.23), offering converging evidence for the reliability of the effect.

exp(deltaBIC/2) (Raftery, 1995; Wagenmakers, 2007). The obtained BF was 7.72, which indicates strong evidence for the presence of the effect.

The proportion of correct responses was .98 and .97 for taboo and neutral words, respectively. Although the pattern is in line with the notion of a taboo-driven facilitation, the analysis of accuracy did not show any effect (χ^2 [1] = 2.003, p = .15, b = 0.39, SE = 0.27, z = 1.42), possibly due to the overall high levels of response accuracy across participants in the task.

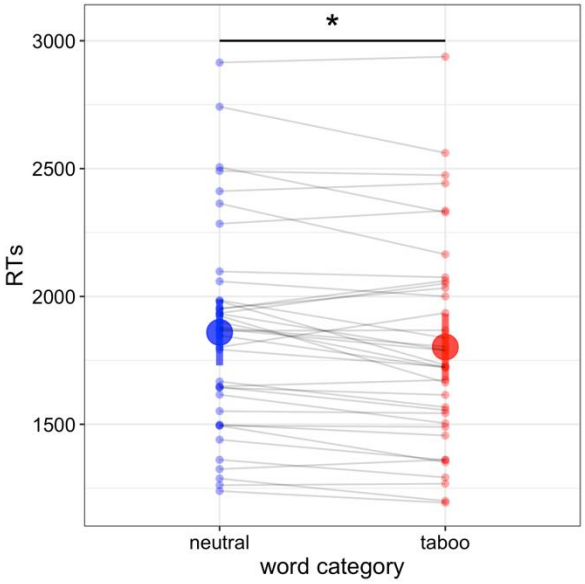

Figure 1. RTs for the two categories of words in Experiment 1. Large points represent by participants mean scores (neutral words: 1860 ms; taboo words: 1802 ms). Error bars indicate 95% confidence intervals across individuals. Small points represent individual means, with lines connecting observations from the same participant.

To investigate habituation, we ran the same analysis as above by also including trial progression and its interaction with word category as fixed factors. The analysis showed no sign of interaction (χ^2 [1] = 1.09, p = .29). The effects of word category (χ^2 [1] = 3.98, p = .04, b = -65.44, SE = 32.72, t = -2.00) and trial progression were significant (χ^2 [1] = 340.44, p < .001, b = -2.43, SE = 0.13, t = -18.76).

Experiment 2

Participants

Forty students from the University of Milano-Bicocca took part in the experiment (22 females, mean age = 22.50; *SD* = 3.55). Participants were all Italian native speakers, and reported normal or corrected-to-normal vision and no history of learning disabilities. They received course credits for their participation, and none of them participated in Experiment 1.

Materials

The same taboo and neutral words of the previous experiment were used. A third set of 70 control words (e.g., vittima, *victim*; henceforth negative words) was also selected from the Italian adaptation of the Affective Norms for English Words database (Montefinese et al., 2014).

Control words were matched with taboo and neutral words on several psycholinguistic variables. Control and taboo words only differed in terms of social appropriateness, and both differed from neutral words on valence and arousal, being overall more negative and more arousing than neutral words (see Table 2).

A new set of 70 neutral (i.e., non-emotional and non-taboo) words were also included as fillers. In this way, half stimuli were emotionally salient (i.e., the taboo and the emotional words) and half stimuli were emotionally neutral (i.e., the neutral and the filler words). This allowed us to keep the proportion of emotional and non-emotional words equal to the one in Experiment 1. Filler words were matched with neutral words on (log) frequency, length, valence, and arousal (all *p*s > .13). As other variables were not controlled, filler words were not included in the analyses.

Table 2

Psycholinguistic variables of the experimental stimuli used in Experiment 2.

Variables	Taboo	Neutral	Control	*p* value Tb vs Neu	*p* value Tb vs Ctr	*p* value Neu vs Ctr
Frequency (log)	8.90	9.63	9.12	.14	.65	.34
Imageability	6.28	6.63	6.51	.13	.13	.63
Concreteness	6.47	6.61	6.23	.53	.26	.10
Valence	3.96	6.07	4.19	<.001	.30	<.001
Arousal	5.18	4.48	5.36	<.001	.14	<.001
N. of Letters	7.23	7.19	7.57	.87	.26	.21
OLD20	2.16	2.02	2.12	.24	.77	.37

Note. Tb = taboo; Neu = neutral; Ctr = control. OLD = orthographic Levenshtein distance and Frequency values (log-transformed) were taken from the SUBTLEX-IT database (Crepaldi et al., 2013). Concreteness, imageability, valence, and arousal scores were taken from ITABOO (Sulpizio et al., 2020) and the Italian adaptation (Montefinese et al., 2014) of the Affective Norms for English Words database (Bradley & Lang, 1999) for taboo and nontaboo words, respectively.

Apparatus and Procedure

The same as in Experiment 1. The only difference was in the duration of the experiment, which was ~30 minutes.

Results

One participant was discarded because of extremely slow response latencies with respect to the sample (participant's RTs were slower than 2.5 *SD* from the sample mean).

The pattern of RTs is represented in Figure 2. The analysis of RTs showed a main effect of word category (χ^2 [1] = 7.51, p = .02)[2]. Pairwise comparisons (Bonferroni corrected) showed that control words were slower than taboo words (b = 78.80, SE = 31.90, t = 2.46, p

[2] The effect of word category is significant also after outlier removal (applied as in Experiment 1, χ^2 [1] = 6.39, p = .04), and when the model is run on log-transformed RTs (and without outlier removal, χ^2 [1] = 6.62, p = .03).

= .04). No difference emerged between control and neutral words (b = 73.21, SE = 31.90, t = 2.29, p = .06), and taboo and neutral words (b = 5.59, SE = 31.90, t = 0.17, p = .98)[3].

Proportion of correct responses were .95, .96, and .97 for control, neutral, and taboo words, respectively, being numerically in line with the pattern of RTs. However, word category failed to reach conventional significance (χ^2 [1] = 5.40, p = .06) in the analyses of accuracy, possibly because of the very high levels of performance consistently displayed across participants in this task.

[3] Since two stimuli (heroin, fat) in the control words category might be considered as taboo (at least according to a more recent database of taboo words, Sulpizio et al., 2024, we re-run the analysis recategorizing these two items as taboo. The overall pattern of the new analyses parallels the original one, although now the contrast between control and neutral words reached conventional significance (main effect of word category: χ^2 [1] = 9.41, p = .009; post-hoc comparisons: control vs taboo words: b = 90.07, SE = 31.80, t = 2.83, p = .01; control vs. neutral words: b = 80.13, SE = 32.00, t = 2.50, p = .03; taboo vs neutral words: t < 1, p >.9).

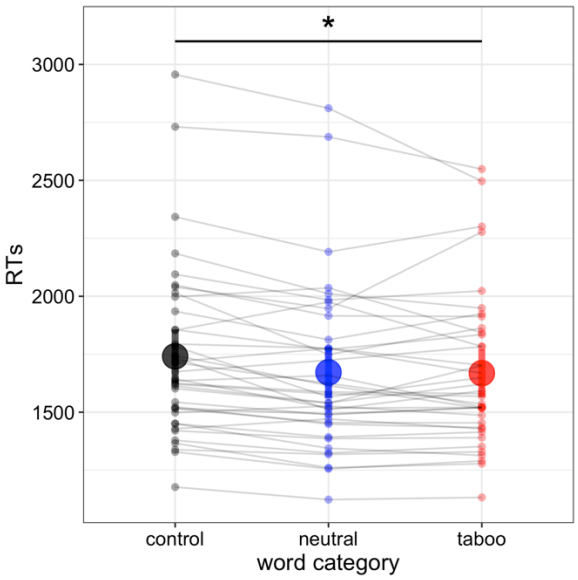

Figure 2. RTs for the three categories of words in Experiment 2. Large points represent by participants mean scores (control words: 1742 ms; neutral words: 1672 ms; taboo words: 1669 ms). Error bars indicate 95% confidence intervals across individuals. Small points represent individual means, with lines connecting observations from the same participant.

In a follow-up analysis, we investigated habituation by running the same analysis as above with the inclusion of trial progression and its interaction with word category as fixed factors. There was no sign of interaction (χ^2 [1] = 0.44, p = .80). The effect of word category was significant, as in the previous analysis (χ^2 [1] = 7.39, p = .02). Also, the effect of trial progression was significant (χ^2 [1] = 605.50, p < .001), with overall RTs decreasing over the course of the experiment. Parameters of the final model are reported in Table S1.

Joint analysis of Experiment 1 and 2.

To further test the asymmetry in the difference between taboo and neutral words emerged across the two experiments, a mixed-effects model on RTs was conducted using word category (taboo vs neutral) and experiment (1 vs 2) as fixed factors, and including by-participants and by-items random intercepts. The analysis showed a significant word category by experiment interaction (χ^2 [1] = 13.73, p < .001). All model parameters are reported in Table S2. Pairwise comparisons (Bonferroni corrected) showed that the difference between

Experiment 1 and 2 was significant for neutral (b = 188, SE = 84.90, t = 2.21, p = .02), but not for taboo words (b = 134, SE = 84.9, t = 1.57, p = .11).

Discussion

The experiments investigated the interplay between semantic properties and perceptual identification of symbolic stimuli. We focused on taboo words and tested whether these stimuli can be recognized earlier than non-taboo words. Using the progressive demasking paradigm, we showed that taboo words were identified faster than neutral words (Experiment 1) and control words (Experiment 2). This effect remained constant throughout the experiment, with no signs of habituation. In Experiment 2, however, taboo and neutral words displayed comparable response latencies, both being identified faster than control words (at least in the post-hoc analyses reported in footnote 2). Speculatively, the lack of difference between taboo and neutral words in Experiment 2 may result from the combined action of two different effects: The taboo advantage, on the one hand, and the higher contextual predictability/contextual priming for neutral words, on the other. These two effects could be simultaneously at work in Experiment 2. Indeed, taboo words in Experiment 2 were as fast as taboo words in Experiment 1 (see Joint Analysis of Experiment 1 and 2 in the manuscript), hinting at a "floor" effect due to the higher attention received by the taboo connotation independently of any list-context effect (taboo words were the 50% of the trials in Experiment 1, but only the 25% of them in Experiment 2). In parallel, responses to neutral words became significantly faster from Experiment 1 to Experiment 2, and this finding could be attributed to increased contextual predictability (with neutral words having a .5 probability of occurrence compared to the .25 for taboo and control words), and greater chance of contextual priming due to category repetition (e.g., Marcel, 1974; Teweedy et al., 1977). Of course, this explanation is entirely post-hoc.

The list composition may also have led the identification system to rely on a guessing mechanism. The use of guessing would have been particularly advantageous for taboo words, given that the number of words falling in this category is relatively small (~200 words in Italian, see Sulpizio et al., 2024), especially when compared to the pool of neutral words. Note, however, that while this guessing hypothesis may account for the results of Experiment 1 (taboo faster than neutral), it seems less plausible for Experiment 2. In Experiment 2, control words, whose affective connotation is comparable to that of taboo words (being, on average, more negative and highly arousing than neutral words), exhibited slower response latencies, despite the comparatively smaller size of this category.

The advantage of taboo connotation over neutral words in Experiment 1 and control words in Experiment 2 aligns with previous research on the temporal dynamics of attention, which consistently reports strong attentional capture phenomena specifically induced by taboo words (e.g., Anderson, 2005; Mathewson et al., 2008). Our results extend this conclusion to a task that requires – but does not directly test – attention allocation, suggesting that the effects of attentional capture extend to word identification. The heightened attention directed toward taboo words may lead to two different outcomes depending on the task-induced processes. When the to-be-performed task mainly relies on visuo-perceptual processes – such as in a perceptual identification task like the progressive demasking paradigm (e.g., Grainger & Segui, 1990) –, the higher attention received by taboo words may produce a processing advantage. However, when the task requires attentional resources to be allocated beyond perceptual identification – such as in the lexical decision task (e.g., Carretié et al., 2008; Scaltritti et al., 2021; Madan et al., 2017) or the Stroop task (Scaltritti et al., 2022) – taboo connotation instead triggers interference effects. This would happen because, in these latter cases, attentional capture would detract resources from attention-demanding task-specific features.

Interestingly, our results suggest that control words fail to trigger attentional capture phenomena comparable to those observed for taboo words, despite having similar affective properties. In Experiment 2, control words displayed the slowest identification time, a pattern consistent with findings from attentional literature showing that attentional capture is specifically triggered by taboo words (Arnell et al., 2007; Mathewson et al., 2008). This relation has been ascribed to stimulus arousal, which is usually higher for taboo than for other categories of words (Arnell et al., 2007; Mathewson et al., 2008). Arousal is known to drive the amount of attention given to a stimulus (e.g., Gronau et al., 2003; Lang et al., 1993; Schimmack & Derryberry, 2005). However, this arousal-based explanation seems at odds with the results of Experiment 2, where taboo and control words were comparable in terms of arousal (and valence). One possible way to reconcile this inconsistency with previous literature is by considering the findings of a recent study by Sulpizio and colleagues (2024), which evaluated several lexico-semantic properties of taboo words. This study showed that *tabooness* – i.e., the extent to which a word is not acceptable in most social situations (Janschewitz, 2008) – is predicted mainly by arousal (and valence), with higher tabooness being associated with higher arousal (and lower valence). Possibly, the effect of arousal reported by previous studies was actually driven by tabooness, for which arousal would have worked as a partial proxy. Indeed, when considering words, a semantic dimension such as tabooness might be even more relevant than arousal (which is known to elicit weaker reactions with respect to pictorial stimuli, e.g., Hinojosa et al., 2009; Kensinger & Schacter, 2006): In fact, tabooness would provide direct information about the appropriateness of the stimulus, thus suggesting whether the specific word can be used in a given context.

One can argue that control words, being comparatively more emotionally charged than neutral ones, may be characterized by a higher tabooness, and should thus resemble taboo words in terms of perceptual facilitation. However, in our perspective taboo and control

words are qualitatively distinct categories. Specifically, taboo words may be marked as socially inappropriate due to their level of offensiveness as well as because of restrictions and sanctions enforced on their use. In contrast, control words, even those that are more affectively connotated (e.g., *funeral* or *victim*), are much less prone to social restrictions, and are arguably not marked as socially inappropriate. Therefore, with respect to this socio-pragmatic dimension, control words would be comparable to neutral ones.

Related to social inappropriateness, tabooness is a semantic property acquired through social and pragmatic linguistic interactions, possibly through classical conditioning. According to some authors, taboo words are originally neutral, gaining their taboo (and emotional) connotation through repeated associations with punishment (Jay et al., 2006; Jay, 2009). This suggests that, under certain conditions, attentional capture may be triggered by socio-culturally determined semantic properties. This represents the intriguing case of a semantic dimension that, although culturally acquired, displays the same effect as phylogenetically-relevant properties (e.g., arousal). In other words, symbolic and non-symbolic stimuli may influence the deployment of attention and modulate subjective awareness in a similar way, although the underlying dynamic would be phylogenetically determined for non-symbolic stimuli, and socially acquired for symbolic ones. This is consistent with previous literature (e.g., Olsson et al., 2005), which suggests that socio-cultural learning of stimulus properties (tabooness of a word, in our case) may form the basis for behavioral responses similar to those triggered by evolutionary-relevant properties found in non-symbolic stimuli (e.g., facial expressions of emotions).

Although the affective connotation can in part be shaped by social and cultural dynamics even in the case of non-symbolic stimuli, we argue that the socio-cultural drive becomes specifically relevant for the taboo dimension, which – as argued above – seems to

be entirely determined by social conventions. We may thus speculate that various social and contextual factors may shape the interaction with the dimension of tabooness.

One socially relevant issue not addressed here, but potentially important for further investigation, is gender differences. A first general issue concerns potential gender differences in the processing of taboo words. This may also further differentiate taboo words from emotional (non-taboo) ones, for which gender differences have occasionally been reported in terms of processing and evaluation (e.g., Haro et al., 2024). Evidence suggests that taboo words are rated similarly by female and male participants across semantic dimensions (e.g., no differences in tabooness or offensiveness; Sulpizio et al., 2024), possibly reflecting a shared societal assessment of the taboo connotation of specific items. To explore this further, we re-analyzed our two experiments, including participants' gender as a covariate. This post-hoc analysis confirmed the main effect of word category (Experiment 1: χ^2 [1] = 3.81, p = .05; Experiment 2: χ^2 [1] = 7.52, p = .02) and no effect of gender (Experiment 1: χ^2 [1] = 0.04, p = .83; Experiment 2: χ^2 [1] = 1.54, p = .21). Given these preliminary results, future research may, however, investigate taboo words with an explicit focus on their social dimension in a more detailed manner, for example, by assessing how taboo words are evaluated and processed by various social groups. In this scenario, gender differences might emerge for items that specifically target the social group of the participant.

The present experiments additionally highlighted that, similarly to previous studies using taboo words in attentional paradigms (Arnell et al., 2007; Mathewson et al., 2008), the taboo connotation advantage during word perceptual identification holds across the entire duration of the experiment, with no hint of habituation. This contrasts with studies showing habituation effects in paradigms where taboo connotation is detrimental (e.g., MacKay et al., 2004; Sulpizio et al., 2022; 2024). The difference, we argue, can be better understood considering task goals and requirements. In the progressive demasking task, the goal is to

identify the stimulus. Since perceptual processing may be enhanced by attention, and since taboo connotation triggers attentional capture, the system may consistently benefit from the full activation of taboo information throughout the experiment, without needing to filter it out through habituation. Differently, in the lexical decision task, the system may dampen taboo-induced attentional capture because the taboo connotation itself is detrimental to the task. In lexical decision, participants must decide on the lexical status of a stimulus, a decision that relies not only on the activation and identification of a specific lexical entry, but also on other parameters including the familiarity and "wordlikeness" of the stimulus (e.g., Balota & Chumbley, 1984; Ratcliff et al., 2004), and the summed activation from orthographically related lexical candidates (e.g., Grainger & Jacobs, 1996). Moreover, the activation of the taboo connotation (as, more in general, semantic information) does not hinge on the recognition of the stimulus as an actual word, as it can also be activated by nonword stimuli (e.g., Bonandrini et al., 2023; Hendrix & Sun, 2021; Sulpizio et al., 2021). Therefore, taboo connotation does not provide information about the lexical status of the stimulus, thus failing to discriminate between the response alternatives and, thus, to directly support the task goal. In this context, the taboo-induced attentional capture would merely drag resources away from the ongoing task. The system would thus intervene, by dampening any activation of the taboo connotation (e.g., Scaltritti et al., 2021; Sulpizio et al., 2022), and, over the course of trials, by filtering the distracting yet task-irrelevant information via habituation (see also Sulpizio et al., 2024).

To conclude, although some caution is warranted, our results show that, under appropriate experimental-context conditions, taboo words may be identified faster than other word categories (i.e., neutral words in Experiment 1 and emotionally-characterized but socially-appropriate words in Experiment 2). This suggests that the relevant semantic properties of purely symbolic stimuli may capture participants' attention, thereby enhancing

perceptual identification. This finding aligns with growing evidence concerning the pervasiveness of semantic processing in visual word recognition (e.g., Balota & Yap, 2016; Carreiras et al., 2014), even within contexts hinging upon earlier stages of perceptual identification and orthographic decoding.

Acknowledgment

We thank Carlo Toneatto for assistance with programming the experiment, and Monica Minelli and Alice Rigola for assistance with data collection. ChatGPT was used to proofread the manuscript.

References

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

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.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B.Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9–32). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human perception and performance, 10*, 340-357.

Balota, D. A., & Yap, M. J. (2006). Attentional control and the flexible lexical processor: Explorations of the magic moment of word recognition. *From inkmarks to ideas: Current issues in lexical processing*. In S. Andrews (Ed.), From inkmarks to ideas: Current issues in lexical processing (pp. 229–258). New York, NY: Psychology Press.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*1, 1-48.

Bertels, J., & Kolinsky, R. (2016). Disentangling fast and slow attentional influences of negative and taboo spoken words in the emotional Stroop paradigm. *Cognition and Emotion*, *30*, 1137-1148.

Bonandrini, R., Amenta, S., Sulpizio, S., Tettamanti, M., Mazzucchelli, A., & Marelli, M. (2023). Form to meaning mapping and the impact of explicit morpheme combination in novel word processing. *Cognitive Psychology*, *145*, 101594.Bradley, M. M., & Lang, P.

J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings (Vol. 30, No. 1, pp. 25-36). Technical report C-1, the center for research in psychophysiology, University of Florida.

Brysbaert, M. (2019). How many participants Do We have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition, 2,* 16.

Brysbaert, M. & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition, 1,* 9.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7*, 308-313.

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in cognitive sciences*, *18*(2), 90-98.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of the orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of experimental psychology: learning, memory, and cognition, 23,* 857.

Carretié, L., Hinojosa, J. A., Albert, J., López-Martín, S., De La Gándara, B. S., Igoa, J. M., & Sotillo, M. (2008). Modulation of ongoing cognitive processes by emotionally intense words. *Psychophysiology, 45,* 188-196.

Cowan, N. (1999). An embedded-processes model of working memory. Models of working memory. *Mechanisms of Active Maintenance and Executive Control*, *20*(506), 1013–1019.

Crepaldi, D., Keuleers, E., Mandera, P., & Brysbaert, M. (2013). SUBTLEX-IT. Retrieved from http://crr.ugent.be/subtlex-it/

Dhooge, E., & Hartsuiker, R. J. (2011). How do speakers resist distraction? Evidence from a taboo picture-word interference task. *Psychological Science, 22*, 855-859.

Duñabeitia, J. A., Avilés, A., & Carreiras, M. (2008). NoA's ark: Influence of the number of associates in visual word recognition. *Psychonomic Bulletin & Review*, *15*, 1072-1077.

Estes, Z., & Adelman, J. S. (2008). Automatic vigilance for nega- tive words in lexical decision and naming: comment on Larsen, Mercer, and Balota (2006). *Emotion*, *8*, 441-444.

Ferrand, L., Brysbaert, M., Keuleers, E., New, B., Bonin, P., Méot, A., ... & Pallier, C. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in psychology*, *2*, 306.

Gayet, S., Paffen, C. L., Belopolsky, A. V., Theeuwes, J., & Van der Stigchel, S. (2016). Visual input signaling threat gains preferential access to awareness in a breaking continuous flash suppression paradigm. *Cognition*, *149*, 77-83.

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics*, *47***,** 191-198.

Gronau, N., Cohen, A., & Ben-Shakhar, G. (2003). Dissociations of personally significant and task-relevant distractors inside and outside the focus of attention: a combined behavioral and psychophysiological study. *Journal of Experimental Psychology: General*, *132*, 512.

Haro, J., Hinojosa, J. A., & Ferré, P. (2024). The role of individual differences in emotional word recognition: Insights from a large-scale lexical decision study. *Behavior Research Methods*, *56*, 8501-8520.

Hendrix, P., & Sun, C. C. (2021). A word or two about nonwords: Frequency, semantic neighborhood density, and orthography-to-semantics consistency effects for nonwords

in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*, 157-183.

Hodsoll, S., Viding, E., & Lavie, N. (2011). Attentional capture by irrelevant emotional distractor faces. *Emotion, 11*, 346-353.

Jay, T. (2009). The utility and ubiquity of taboo words. *Perspectives of Psychological Science, 4*, 153-161.

Jay, T., King, K., & Duncan, T. (2006). Memories of punishment for cursing. *Sex Roles*, *55*, 123-133.

Janschewitz, K. (2008). Taboo, emotionally valenced, and emotionally neutral word norms. *Behavior research methods, 40*, 1065-1074.

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*, 473-481.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7*, 12-18.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*, 261-273.

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.

Madan, C. R., Shafer, A. T., Chan, M., & Singhal, A. (2017). Shock and awe: Distinct effects of taboo words on lexical decision and free recall. *Quarterly Journal of Experimental Psychology*, *70*, 793-810.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of physiology-Paris*, *102*, 59-70.

Marcel, T., & Forrin, B. (1974). Naming latency and the repetition of stimulus categories. *Journal of Experimental Psychology, 103*, 450-460.

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.

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the affective norms for English words (ANEW) for Italian. *Behavior Research Methods, 46*, 887-903.

Moray, N., Fitter, M., Ostry, D., Favreau, D., & Nagy, V. (1976). Attention to pure tones. *Quarterly Journal of Experimental Psychology, 28*, 271-283.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological review*, *108*, 483-522.

Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The role of social groups in the persistence of learned fear. *Science*, *309*, 785-787.

Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 261-282.

Pulvermuller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain. *Brain and language, 110*, 81e94.

Rabagliati, H., Robertson, A., & Carmel, D. (2018). The importance of awareness for understanding language. *Journal of Experimental Psychology: General, 147*, 190-208.

R Core Team. (2015). R: A language and environment for statistical comput- ing. R Foundation for Statistical Computing. http://www.R-project.org/

Sergent, C., Wyart, V., Babo-Rebelo, M., Cohen, L., Naccache, L., & Tallon-Baudry, C. (2013). Cueing attention after the stimulus is gone can retrospectively trigger conscious perception. *Current Biology, 23*, 150-155.

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences, 109*, 19614-19619.

Scaltritti, M., Job, R., & Sulpizio, S. (2021). Selective Suppression of Taboo Information in Visual Word Recognition: Evidence for Cognitive Control on Semantics. *Journal of Experimental Psychology: Human Perception & Performance, 47*, 934-945.

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion* (pp. 92-120). Oxford, England: Oxford University Press.

Schimmack, U., & Derryberry, D. (2005). Attentional Interference Effects of Emotional Pictures: Threat, Negativity, or Arousal? *Emotion, 5*, 55-66.

Schreuder, R., and Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory & Language, 37*, 118-139.

Schubring, D., & Schupp, H. T. (2019). Affective picture processing: Alpha-and lower beta-band desynchronization reflects emotional arousal. *Psychophysiology, 56*, e13386.

Sulpizio, S., Arcara, G., Lago, S., Marelli, M., & Amenta, S. (2022). Very early and late form-to-meaning computations during visual word recognition as revealed by electrophysiology. *Cortex, 157*, 167-193.

Sulpizio, S., Günther, F., Badan, L., Basclain, B., Brysbaert, M., Chan, Y. L.,..., & Marelli, M. (2024). Taboo language across the globe: A multi-lab study. *Behavior Research Methods*.

Sulpizio, S., Job, R., Leoni, P., & Scaltritti, M., (2022). Prepotent task-irrelevant semantic information is dampened by domain-specific control mechanisms during visual word recognition. *Quarterly Journal of Experimental Psychology, 75*, 390-405.

Sulpizio, S., Pennucci, E., & Job, R. (2021). The impact of emotional content on pseudoword recognition. *Psychological Research, 85*, 2980-2996.

Sulpizio, S., Scaltritti, M., & Spinelli, G. (2024). Fast habituation to semantic interference generated by taboo connotation in reading aloud. *Cognition & Emotion*.

Sulpizio, S., Toti, M., Del Maschio, N., Costa, A., Fedeli, D., Job, R., & Abutalebi, J. (2019). Are you really cursing? Neural processing of taboo words in native and foreign language**.** *Brain & Language, 194*, 84-92.

Sulpizio, S., Vassallo, E., Job, R., & Abutalebi, J. (2020). ITABÙ: Dati preliminari di un database delle parole tabù per l'italiano [ITABOO: Preliminary data for an Italian database for taboo words]. *Giornale Italiano di Psicologia, 2,* 599-614.

Tweedy, J. R., Lapinski, R. H., & Schvaneveldt, R. W. (1977). Semantic-context effects on word recognition: Influence of varying the proportion of items presented in an appropriate context. *Memory & Cognition, 5*, 84-89.

Whiting, C. M., Cowley, R. G., & Bozic, M. (2017). The Role of Semantic Context in Early Morphological Processing. *Frontiers in Psychology, 8*.

Wickham, H. (2016). Ggplot2: Elegant graphics for data analysis. Springer.

Williams, L. J., & Evans, J. R. (1980). Evidence for perceptual defense using a lexical decision task. *Perceptual and Motor Skills, 50*, 195-198.

Yang, Y. H., & Yeh, S. L. (2011). Accessing the meaning of invisible words. *Consciousness and Cognition, 20,* 223-233.

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion, 7*, 882-886.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic bulletin & review, 15,* 971-979.

Supplemental Materials

Table S1.
Parameters of the model of the analysis testing habituation in Experiment 2.

Random effects	Variance	SD	
Participant	118039	343.6	
Item	31935	178.70	
Residual	137504	370.80	
Fixed Effects	b	SE	t
Intercept	1938.68	59.94	32.34
Word category (neutral)	-72.81	31.89	-2.28
Word category (taboo)	-77.65	31.88	-2.43
Trial progression	-1.31	0.05	-25.09

Table S2.
Parameters of the model of the joint analysis of Experiment 1 and 2.

Random effects	Variance	SD	
Participant	136745	369.80	
Item	30884	175.70	
Residual	137646	371.00	
Fixed Effects	b	SE	t
Intercept	1873.26	63.94	29.29
Word category (taboo)	-70.51	31.36	-2.24
Experiment (2)	-187.90	84.92	-2.21
Word category x Experiment	53.81	14.52	3.70