Affective Valence of Words Differentially Affects Visual and Auditory Word Recognition

Chuanji Gao[a], Svetlana V. Shinkareva[b], and Marius V. Peelen[a]

[a]*Donders Institute for Brain, Cognition and Behaviour, Radboud University, 6525 HR, Nijmegen, The Netherlands*

[b]*Department of Psychology, Institute for Mind and Brain, University of South Carolina, Columbia, SC 29201, USA*

Correspondence should be addressed to Chuanji Gao, Donders Institute for Brain, Cognition and Behaviour, Radboud University, 6525 HR, Nijmegen, The Netherlands, email: chuanji.gao@donders.ru.nl.

Declarations of interest: none.

Main text word count (excluding title, references, affiliations, acknowledgements, figures and figure legends): 8678.

Number of references: 92.

Acknowledgement: This project was enabled by a Radboud Excellence fellowship from Radboud University in Nijmegen, the Netherlands, and has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 725970). We thank Rutvik Desai for help in putting together the databases. We thank Surya Gayet, Roel Willems, Genevieve Quek, Victor Kuperman, and Marc Brysbaert for helpful comments on the manuscript.

Data availability statement: The data, code and modeling output files for the present study has been made available at Open Science Framework (https://osf.io/upnav/).

Abstract

Recognizing written or spoken words involves a sequence of processing stages, transforming sensory features into lexical-semantic representations. While the later processing stages are common across modalities, the initial stages are modality-specific. In the visual modality, previous studies have shown that words with positive valence are recognized faster than neutral words. Here, we examined whether the effects of valence on word recognition are specific to the visual modality or are common across visual and auditory modalities. To address this question, we analyzed multiple large databases of visual and auditory lexical decision tasks, relating the valence of words to lexical decision times while controlling for a large number of variables, including arousal and frequency. We found that valence differentially influenced visual and auditory word recognition. Valence had an asymmetric effect on visual lexical decision times, primarily speeding up recognition of positive words. By contrast, valence had a symmetric effect on auditory lexical decision times, with both negative and positive words speeding up word recognition relative to neutral words. The modality-specificity of valence effects were consistent across databases and were observed when the same set of words were compared across modalities. We interpret these findings as indicating that valence influences word recognition partly at the sensory-perceptual stage. We relate these effects to the effects of positive (reward) and negative (punishment) reinforcers on perception.

Keywords: emotion, valence, word recognition, visual, auditory

Affective Valence of Words Differentially Affects Visual and Auditory Word Recognition

The capacity to recognize written and spoken words is an essential human cognitive skill. One effective way to uncover the cognitive processes of word recognition is to examine the time it takes to distinguish words from non-words (Balota & Chumbley, 1984; Balota et al., 2006; Ferrand et al., 2018; Goldinger, 1996; Grainger, 1990; Grainger & Jacobs, 1996; Jacobs & Grainger, 1994; Meyer & Schvaneveldt, 1971; Yap & Balota, 2009; Ziegler et al., 2003). In a *lexical decision task*, participants indicate whether a written or spoken item is a word or a non-word as fast and as accurately as possible. The resulting lexical decision time has been used as an index of lexical access, providing insight into the functional and computational mechanisms of word recognition (Grainger & Jacobs, 1996; Jacobs & Grainger, 1994; McClelland & Rumelhart, 1981; Ratcliff et al., 2004). Initial studies primarily investigated the effect of psycholinguistic variables such as word frequency, word length, age of acquisition, and concreteness on lexical decision time (Carreiras et al., 1997; Gerhand & Barry, 1999; Gordon, 1985; Hudson & Bergman, 1985; Jessen et al., 2000).

Intriguingly, recent studies investigating visual word recognition showed that the affective valence of words also affects lexical decision time (Estes & Adelman, 2008b; Kissler & Koessler, 2011; Kousta et al., 2009; Kuperman et al., 2014; Larsen et al., 2006; Larsen et al., 2008; Vinson et al., 2014; Yap & Seow, 2014). Because lexical decision time reflects the outcome of a sequence of cognitive processes - from sensory processing to semantic memory - it is unclear at what processing stage(s) valence affects word recognition. In the present study, we addressed this question by directly comparing the effect of valence on visual and auditory word recognition. If valence differentially affects visual and auditory lexical decision times, this would provide evidence for an influence of valence on sensory-perceptual components of word processing, considering that higher-level processing stages are shared between visual and auditory word recognition (Caramazza & Hillis, 1990; Deniz et al., 2019; McClelland & Rumelhart, 1981; Rogowsky et al., 2016; Rumelhart & McClelland, 1982; Segaert et al., 2012).

The specific influence of valence on visual word recognition is still under debate. Some studies found a categorical relationship between valence and visual word recognition (**Figure 1a**), in which reaction times are relatively slow for negative valence and relatively fast for positive valence (Estes & Adelman, 2008a, 2008b) but without much difference within each valence category. Other studies instead found an inverted-U function (**Figure 1b**), wherein positive and negative words have faster reaction times compared to neutral words, and there is no difference between positive and negative words (Kousta et al., 2009; Vinson et al., 2014; Yap & Seow, 2014). Finally, there is also evidence for a monotonic relationship between valence and visual word recognition (Kuperman et al., 2014), with slower reaction times for more negative and faster reaction times for more positive words (**Figure 1c**). These differences between studies may be due to the specific word set studied and, particularly, whether or not additional control variables were included in the model (e.g., frequency and valence-by-frequency interaction; Kuperman et al., 2014).

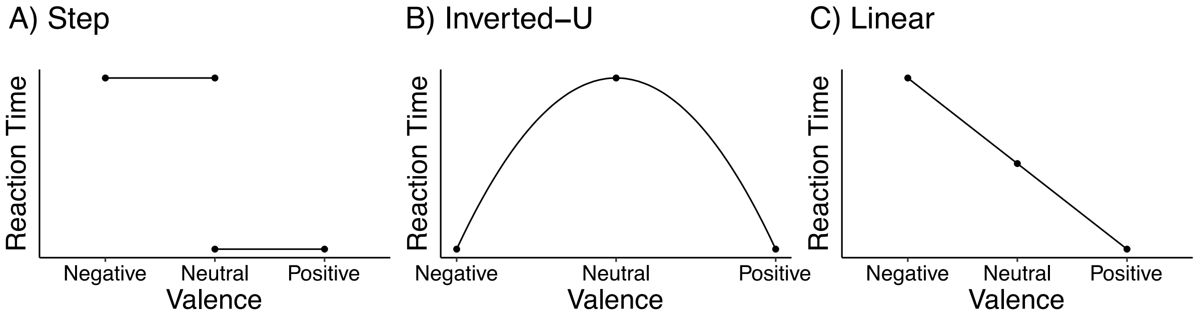

Figure 1. Summary of possible relations between valence and lexical decision time. A) step function; B) inverted-U function; and C) linear function.

Though several studies have examined the effect of valence on visual word recognition, the effect of valence on auditory word recognition has received far less attention. One previous study found significant effects for both linear and quadratic valence terms on auditory lexical decision time using a sample of 514 words (Goh et al., 2016). However, the study did not control for variables such as age-of-acquisition, contextual diversity, initial phoneme, and valence-by-frequency interaction. Furthermore, results were not compared with the corresponding visual lexical decision times. Therefore, the specific influence of valence on auditory lexical decision, and how this compares with visual lexical decision, warrants further investigation.

We examine the effects of valence on visual and auditory lexical decision times in three investigations. First, we analyzed a large sample of common words to examine the effects of valence on visual and auditory lexical decision times, in which visual lexical decision times were drawn from the widely-used English Lexicon Project (i.e., ELP; Balota et al., 2007) and auditory lexical decision times were taken from the largest auditory lexical decision database available (Massive Auditory Lexical Decision, i.e., MALD) (Tucker et al., 2019). Having derived a custom set of words for which both visual and auditory lexical decision data were available, we set about obtaining a wide array of relevant covariates for this stimulus list, including frequency, contextual diversity, age-of-acquisition, and orthographic neighborhood. Based on previous studies (Kahan & Hely, 2008; Kuperman et al., 2014; Scott et al., 2009; Sheikh & Titone, 2013), interactions of valence and arousal with frequency were also considered.

Second, we examined whether or not the effects of the first analysis would generalize to other large databases. Specifically, we investigated visual and auditory lexical decision times from five databases: ELP, British Lexicon Project (i.e., BLP; Keuleers et al., 2012) and English Crowdsourcing Project (i.e., ECP; Mandera et al., 2019) for visual lexical decision; and MALD and Auditory English Lexicon Project (i.e., AELP; Goh et al., 2020) for auditory lexical decision.

Finally, we examined the effects of valence on lexical decision when highly reliable visual and auditory lexical decision times were obtained by averaging visual lexical decision times across ELP, ECP, BLP, and averaging auditory lexical decision times across MALD and AELP, though for a more limited set of words (considering that these had to be shared across all databases).

All three analyses revealed that valence differentially affects visual and auditory word recognition. The effect of valence on visual word recognition was asymmetric, in which positive valence facilitated word recognition. By contrast, the effect of valence on auditory word recognition was symmetric, with both positive and negative valence facilitating word recognition relative to neutral words.

Method

Data

For Analysis 1, we first created a list of 9572 words that are found both in the ELP database of visually presented words, and the MALD database of auditorily presented words. From these two databases, we retrieved the response times to the visually and auditorily presented words (i.e., dependent variables), and for each word we retrieved a number of predictor variables, including valence, arousal, word frequency, etc. (see the description in the next sections). Note that the number of 9572 words is maximal for ELP and MALD with the inclusion of all other lexical control variables (**Figure 2A**).

For Analysis 2, Data were 4149 words including all of the variables of Analysis 1: Emotion predictors, lexical control predictors and dependent variables. Different from Analysis 1, three additional dependent variables were also included: mean visual lexical decision times from ECP and BLP; and mean auditory lexical decision times from AELP. Thus, there were lexical decision times from five different databases (ELP, ECP, BLP, MALD, and AELP) along with a common set of emotion and control variables (**Figure 2B**).

For Analysis 3, Data were the same 4149 words as used in Analysis 2 with emotion predictors and lexical control predictors. Different from Analysis 2, visual lexical decision times from ELP, ECP and BLP were averaged per word as a common visual lexical decision time variable; and auditory lexical decision times from MALD and AELP were averaged as a common auditory lexical decision time variable (**Figure 2B**). The research ethics committee approval is not applicable for the current study given it analyzed public available datasets and do not involve any data collection.

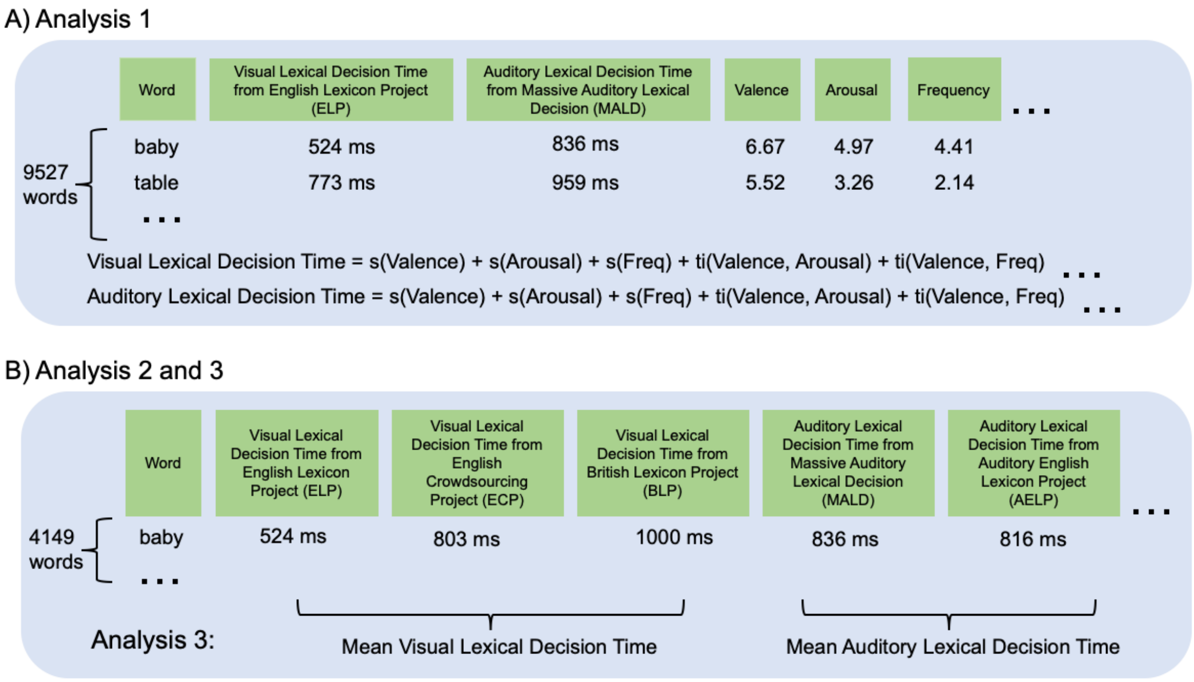

Figure 2. Schematic illustration of the data used for A) Analysis 1, B) Analysis 2 and 3.

Dependent variables

For Analysis 1, mean visual lexical decision time were retrieved from ELP (Balota et al., 2007). Mean auditory lexical decision time were retrieved from MALD (Tucker et al., 2019). For Analysis 2 and 3, mean visual lexical decision times were retrieved from ELP (Balota et al., 2007), ECP (Mandera et al., 2019) and BLP (Keuleers et al., 2012); mean auditory lexical decision times were retrieved from MALD (Tucker et al., 2019) and AELP (Goh et al., 2020).

Emotion predictors

Valence and arousal ratings were drawn from Warriner et al. (2013). Participants were asked to rate how they felt while reading each word on scales ranging from 1 (happy) to 9 (unhappy) and 1 (alert) to 9 (calm).

Lexical control predictors

Word frequencies (Freq) and contextual diversity (CD) were retrieved from Brysbaert and New (2009). Written frequencies (Freq_CobW) and spoken frequencies (Freq_CobS) were retrieved from CELEX (http://celex.mpi.nl/) (Baayen et al., 1996). Age-of-acquisition (AoA) was drawn from Kuperman et al. (2012). Length-related variables including number of letters (NLetters), number of phonemes (NPhon), number of syllables (NSyll), and number of morphemes (NMorph) were retrieved from the ELP (Balota et al., 2007). Lexical density-related variables including orthographic (Ortho_N), phonological (Phono_N), phonographic (Phonographic_N) neighborhoods, orthographic (OLD) and phonological (PLD) Levenshtein distance (The mean of the closest 20 Levenshtein distance neighbors for the orthography or phonology) were also retrieved from ELP. Bigram frequency (BG) was drawn from ELP. Concreteness was drawn from Brysbaert et al. (2014). Orthographic uniqueness point (OUP), Phonological uniqueness point (PUP) and duration of each time were retrieved from MALD database (Tucker et al., 2019). Dominant part of speech (DPoS) was retrieved from Brysbaert et al. (2012). Initial phoneme was obtained from the CMU Pronouncing Dictionary (http://www.speech.cs.cmu.edu/cgi-bin/cmudict) (**Table 1**).

Table 1. Lexical Control Variables

Lexical control variable	Abbreviation	Reference
Word frequency	Freq	Brysbaert and New (2009)
Contextual diversity	CD	Brysbaert and New (2009)
Written frequency	Freq_CobW	Baayen et al. (1996)
Spoken frequency	Freq_CobS	Baayen et al. (1996)
Age-of-acquisition	AoA	Kuperman et al. (2012)
Number of letters	NLetters	Balota et al. (2007)
Number of phonemes	NPhon	Balota et al. (2007)
Number of syllables	NSyll	Balota et al. (2007)
Number of morphemes	NMorph	Balota et al. (2007)
Orthographic neighborhood	Ortho_N	Balota et al. (2007)
Phonological neighborhood	Phono_N	Balota et al. (2007)
Phonographic neighborhood	Phonographic_N	Balota et al. (2007)
Orthographic Levenshtein distance	OLD	Balota et al. (2007)
Phonological Levenshtein distance	PLD	Balota et al. (2007)
Bigram frequency	BG	Balota et al. (2007)
Orthographic uniqueness point	OUP	Tucker et al. (2019)

Phonological uniqueness point	PUP	Tucker et al. (2019)
Concreteness	Concreteness	Brysbaert et al. (2014)
Dominant part of speech	DPoS	Brysbaert et al. (2012)
Initial phoneme		CMU Pronouncing Dictionary

Data Preprocessing

Coding of initial phoneme

For initial phoneme, based on phonemic-level transcription in international phonetic alphabet format from CMU Pronouncing Dictionary, each word was coded dichotomously (1 or 0) according to 14 categories: bilabial, labiodental, interdental, alveolar, palatal, velar, glottal, stop, fricative, affricate, nasal, liquid, glide, and voiced (see Balota et al., 2004; Spieler & Balota, 1997; Treiman et al., 1995 for similar coding). In this coding, 1 denotes the presence of a feature and 0 denotes the absence of a feature.

Transformation of variables

The included Freq, Freq_CobW, Freq_CobS and CD variables were $\log_{10}$-transformed. Given that distributions of reaction times are typically skewed, we $\log_{10}$-transformed visual and auditory lexical decision times. All quantitative variables, including dependent variables, were standardized to have a mean of 0 and standard deviation of 1 before being included into statistical analyses.

Treatment of multicollinearity

To reduce multicollinearity, first, principal component analysis (PCA) was conducted on 12 length-related and lexical density-related variables: Nletters, NPhon, NSyll, NMorph, Duration, PLD, OLD, OUP, PUP, Ortho_N, Phono_N, and Phonographic_N. For Analysis 1, the eigenvalues for the first two components were both larger than 1 and explained 83% of the variance. For Analyses 2 and 3, the first two components from the PCA analysis included in the model explained 85% of the variance. Thus, the first two components (labeled as PC1 and PC2) were included into statistical analyses as predictors.

Second, given that CD and AoA are typically correlated with frequency, the variance of Freq_CobW (or Freq) was partialled out from AoA and CD for the visual lexical decision analysis and the variance of Freq_CobS (or Freq) was partialled out from AoA and CD for the auditory lexical decision analysis. The residuals (labeled as rAoA and rCD) were used in further statistical analyses. A residualization approach was used here to be able to model the interaction between frequency and valence, reported in previous literature (Kahan & Hely, 2008; Kuperman et al., 2014; Scott et al., 2009; Sheikh & Titone, 2013), without the variance of frequency being taken away due to the inclusion of CD and AoA. Note that we were not interested in the unique explanatory power of frequency beyond that of CD and AoA (see Wurm & Fisicaro, 2014 for a cautionary note on using the residualization approach).

Statistical Analyses

Before running statistical analyses, we checked the description of quantitative variables such as correlations between all quantitative predictor variables, and zero-order correlations between each quantitative predictor and each of the dependent variables. We observed that dependent variables (lexical decision times) were skewed as expected. Variables were clustered as meaningful clusters such as length-related cluster, neighborhood measures and frequency measures (**Supplemental Figures 1-3**). To determine potential nonlinear relationships between the valence predictor and dependent variables, generalized additive modeling (GAM) approaches (Hastie & Tibshirani, 1990; Wood, 2017) were implemented with *mgcv* R package (Wood & Wood, 2015; see https://cran.r-project.org/web/packages/mgcv/index.html). We used the general purpose thin plate regression splines (TPRS) as the smoother. Each smoother is a sum of K basis functions multiplied by coefficients. In *mgcv* package, the maximum complexity of each smoother is determined by K. We used the default K = 10 for TPRS (as the effects we examined are not expected to be as complex as having 9 knots) and checked the effective degrees of freedom in the model output to see if K needed to increase (i.e., whether effective degrees of freedom is close to 9). The order of penalized derivatives was set to be M = 2, i.e., penalty is proportionate to the integral of squared second derivatives, which is a typical setting for TPRS. We chose REML as the smoothing parameter estimation method and newton as the numerical optimizer. The double penalty approach was used to produce an extra penalty so that a smooth component can be completely removed (Marra & Wood, 2011), which means that the effective degrees of freedom would be equal to or less than 1 if the model penalized the smooth term to a simple linear relationship. In this way, we do not need to re-enter smoothers into models as linear predictors. We included corresponding frequency measures for visual ($Freq_{CobW}$) and auditory ($Freq_{CobS}$) lexical decision based on the hypothesis that written frequency may match well with visual lexical decision and spoken frequency may match well with auditory lexical decision. We have also performed additional analyses with the same general word frequency (Freq) included for both visual and auditory models, which showed similar findings as when including corresponding written and spoken frequencies for the visual and auditory models, respectively (Supplemental material).

The model structure for the visual lexical decision was:

$$LDT_V = s(Valence) + s(Arousal) + s(Freq_{CobW}) + ti(Valence, Arousal) + ti\ (Valence, Freq_{CobW}) + ti(Arousal, Freq) + s(BG) + s(Concreteness) + s(rCD) + s(rAoA) + s(PC1) + s(PC2) + DPoS + Initial\ Phoneme + \varepsilon$$
(Equation 1)

wherein LDT_V indicates visual lexical decision time, s() refers to the smoother, ε refers to the residuals, and ti() refers to the tensor product interaction with TPRS as basis function.

The model structure for the auditory lexical decision was:

$$LDT_A = s(Valence) + s(Arousal) + s(Freq_{CobS}) + ti(Valence, Arousal) + ti(Valence, Freq_{CobS}) + ti(Arousal, Freq) + s(BG) + s(Concreteness) + s(rCD) + s(rAoA) + s(PC1) + s(PC2) + DPoS + Initial\ Phoneme + \varepsilon$$
(Equation 2)

wherein LDT_A indicates auditory lexical decision time, s() refers to the smoother and ε refers to residuals, and ti() refers to the tensor product interaction with TPRS as basis function. Note that ti() rather than te() was used here because ti() produces a tensor product interaction (i.e., pure smooth 'interaction' component with main effects excluded) while te() produces a full tensor product smooth (See pages 232, 325 in Wood, 2017; see also https://cran.r-project.org/web/packages/mgcv/index.html).

Specifically, for Analysis 1, GAM modelling approaches were performed with lexical decision times from ELP and MALD of a common set of words. For Analysis 2, the same GAM modelling approach was used as in Analysis 1, with five different dependent variables of a common set of words: visual lexical decision time from ELP, visual lexical decision time from ECP, visual lexical decision time from BLP, auditory lexical decision time from MALD, and auditory lexical decision time from AELP. For Analysis 3, the same GAM modelling approach was used as in Analyses 1 and 2 for visual (mean of ELP, ECP and BLP) and auditory (mean of MALD and AELP) lexical decision times. The data, code and modeling output files for the present study has been made available at Open Science Framework (https://osf.io/upnav/).

Results and Discussion

Analysis 1: ELP and MALD

The Effect of Valence on Visual Lexical Decision Time

Generalized additive modeling (GAM) was used to examine the (possibly nonlinear) relation between valence and word recognition (**Figure 1**). We found that valence significantly predicted visual lexical decision time (**Table 2**). The effective degrees of freedom provide information about the number of turning points used to model the relationship between the predictor and the dependent variable with smooth penalties applied. The effective degrees of freedom for valence were close to 1 (i.e., 1.140), suggesting that the effect of valence on visual lexical decision time was close to a simple linear relationship. **Figure 3A** showed that the influence of valence on visual lexical decision time was best described by a monotonic function, in which the rate of change remains constant across the whole valence range, in which the more positive, the faster the lexical decision. This result replicates the findings of Kuperman et al. (2014) despite slight changes in the model used (e.g., addition of PUP and OUP variables and different coding of initial phoneme).

Though our main hypotheses are about the relation between valence and word recognition, the effects of other predictors are also reported for completeness. The effect of arousal on visual lexical decision time was not significant (**Table 2**). As expected, the effect of written frequency on visual lexical decision time showed that the larger the frequency, the shorter the reaction time (**Supplemental Figure 4A**). We also found interactions between valence and written frequency: the effect of valence was largest when frequency was low (**Supplemental Figure 4A**). Functional relationships of other significant predictors with visual lexical decision time are also shown in **Supplemental Figure 4A**. Finally, comparable effects of valence on visual lexical decision time were found when general frequency (Brysbaert & New, 2009) rather than written frequency was modeled (**Supplemental Figure 5**).

Table 2. GAM results for visual lexical decision time.

Smooth terms	edf	Ref.df	F	*p*
Valence	1.140	9	4.409	p < 0.001
Arousal	1.071	9	0.234	p = 0.128
Freq_CobW	6.693	9	204.8	p < 0.001
BG	2.613	9	5.898	p < 0.001
Concreteness	2.891	9	2.320	p < 0.001
rCD	4.733	9	100.6	p < 0.001
rAoA	4.086	9	47.60	p < 0.001
PC1	6.561	9	185.3	p < 0.001

PC2	1.684	9	8.874	p < 0.001
Valence × Arousal	0.906	80	0.024	p = 0.099
Valence × Freq_CobW	6.472	77	0.252	p < 0.001
Arousal × Freq_CobW	2.042	76	0.056	p = 0.057

Note. edf refers to effective degrees of freedom. Ref.df refers to reference degrees of freedom. The *F* is a Wald-type test statistic based on Nychka's across-the-function interpretation of Bayesian credible intervals for the spline (Wood, 2013).

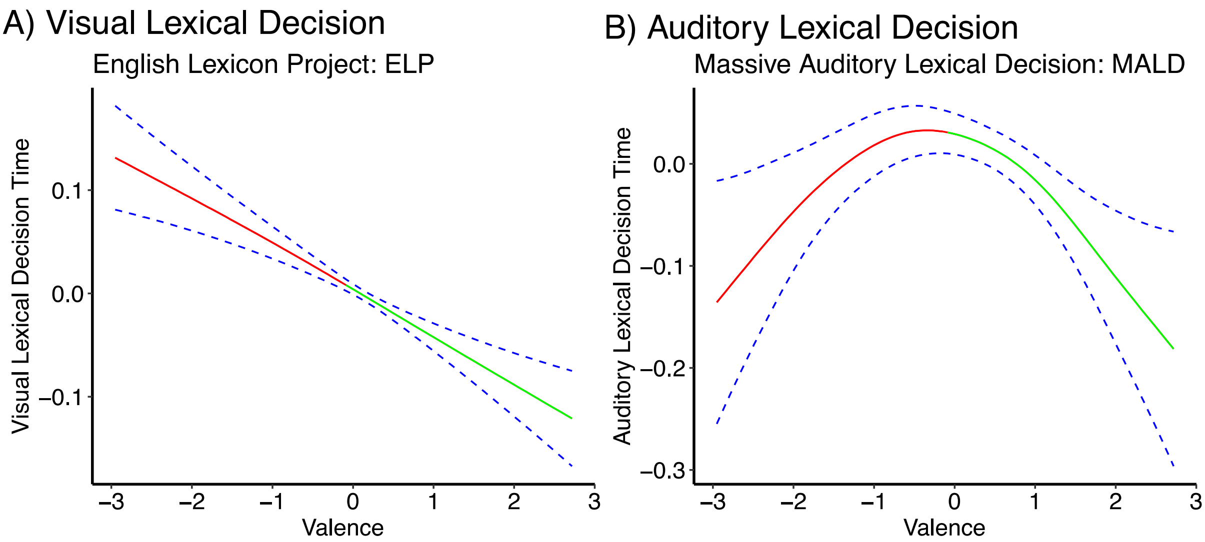

Figure 3. GAM functions describing the relationship between valence in Analysis 1 with A) visual lexical decision time for English Lexicon Project (ELP) and B) auditory lexical decision time for Massive Auditory Lexical Decision (MALD). Green = Positive valence; red = negative valence. Dotted blue lines show 95% confidence intervals for GAM functions.

The Effect of Valence on Auditory Lexical Decision Time

Using the same set of words, we next examined the effect of valence on auditory lexical decision time. We found that valence significantly predicted auditory lexical decision time (**Table 3**). Strikingly, in contrast to the linear function found for visual lexical decision time, the effect of valence on auditory lexical decision time followed an inverted-U function (edf = 2.496), wherein positive and negative valences are nearly symmetric (**Figure 3B**).

The effect of arousal on auditory lexical decision time was significant, in which descriptively there was a slower rate of change during low arousal and a faster rate of change for high arousal (**Supplemental Figure 4B**). The auditory lexical decision time decreased with increased spoken frequency, as expected. We also found interactions between arousal and spoken frequency: the effect of arousal was largest when spoken frequency was low (**Supplemental Figure 4B**). Functional relationships of other significant predictors with auditory lexical decision are shown in **Supplemental Figure 4B**. Finally, comparable effects of valence on auditory lexical decision were found when general frequency (Brysbaert & New, 2009) rather than spoken frequency was modeled (**Supplemental Figure 5**).

Table 3. GAM results for auditory lexical decision.

Smooth terms	edf	Ref.df	F	p
Valence	2.496	9	1.570	$p < 0.001$
Arousal	1.698	9	0.760	$p = 0.011$
Freq	2.081	9	15.44	$p < 0.001$
BG	0.007	9	$F < 0.001$	$p = 0.755$
Concreteness	0.703	9	0.261	$p = 0.066$
rCD	2.798	9	5.499	$p < 0.001$
rAoA	0.993	9	11.57	$p < 0.001$
PC1	4.136	9	15.24	$p < 0.001$
PC2	1.527	9	2.956	$p < 0.001$
Valence × Arousal	0.378	79	0.006	$p = 0.240$
Valence × Freq_CobS	1.711	77	0.032	$p = 0.172$
Arousal × Freq_CobS	3.716	77	0.119	$p = 0.013$

Note. edf refers to effective degrees of freedom. Ref.df refers to reference degrees of freedom. The *F* is a Wald-type test statistic based on Nychka's across-the-function interpretation of Bayesian credible intervals for the spline (Wood, 2013).

Comparing the Effects of Valence on Visual and Auditory Lexical Decision Times

The above analyses for the effects of valence on visual and auditory lexical decision times showed differential functional forms, in which the auditory functional form looks symmetrical while the visual does not. To statistically contrast the influence of valence on visual versus auditory lexical decision, we performed two analyses with the factor *Modality* (visual, auditory) included. First, to evaluate whether the functional forms are different for the visual versus the auditory modality, we examined if there was an interaction between *Modality* and *Valence* by including both variables into one overall model. Note that general frequency rather than written or spoken frequency was used in this analysis to have a common frequency variable across visual and auditory lexical decision times. The model structure for this analysis was:

$$LDT = Modality + s(Valence, by = Modality) + \mathrm{s(Arousal)} + s(Freq) + ti(Valence, Arousal) + ti(Valence, Freq) + ti(Arousal, Freq) + s(BG) + s(Concreteness) + s(rCD) + s(rAoA) + s(PC1) + s(PC2) + DPoS + Initial\ Phoneme + \varepsilon \quad \text{(Equation 3)}$$

wherein s(Valence, by = Modality) indicates an interaction between valence and modality.

A significant interaction between valence and modality was found ($F_{edf=1.98,\ Ref.df=9} = 0.990$, $p = 0.005$), indicating a differential influence of valence on visual and auditory lexical decision times.

Second, we performed an analysis specifically comparing the effect of valence on visual versus auditory lexical decision in terms of symmetry. In this analysis, *Valence* was changed from a continuous variable to a binary variable (positive, negative) using median split.

The model structure for this analysis was:

$$LDT = Valence + Modality + Valence \times \text{Modality} + \text{s(Arousal)} + s(Freq) + s(Arousal, by = Valence) + s(Freq, by = Valence) + ti(Arousal, Freq) + s(BG) + s(Concreteness) + s(rCD) + s(rAoA) + s(PC1) + s(PC2) + DPoS + Initial\ Phoneme + \varepsilon \text{ (Equation 4)}$$

wherein s(Arousal, by = Valence) indicates an interaction between arousal and valence. Follow-up analyses were performed for visual and auditory with the following model structures:

$$LDT_V = Valence + s(Arousal) + s(Freq) + s(Arousal, by = Valence) + s(Freq, by = Valence) + ti(Arousal, Freq) + s(BG) + s(Concreteness) + s(rCD) + s(rAoA) + s(PC1) + s(PC2) + DPoS + Initial\ Phoneme + \varepsilon \text{ (Equation 5)}$$

$$LDT_A = Valence + s(Arousal) + s(Freq) + s(Arousal, by = Valence) + s(Freq, by = Valence) + ti(Arousal, Freq) + s(BG) + s(Concreteness) + s(rCD) + s(rAoA) + s(PC1) + s(PC2) + DPoS + Initial\ Phoneme + \varepsilon \text{ (Equation 6)}$$

A significant interaction between valence and modality was found (t = 7.117, p < 0.001), indicating a differential influence of valence on visual and auditory lexical decision times. The valence effect was significant for visual (t = -6.043, p < 0.001) but not for auditory (t = -1.578, p = 0.115), suggesting that the effects of valence on visual lexical decision were asymmetric (i.e., negativity slow-down effects) while the effects of valence on auditory lexical decision were symmetric.

Effects for Distinct Emotion Categories

The above analyses showed that the effects of positive valence were similar across modalities while the effects of negative valence were different. To test whether the negative valence difference was driven by specific negative emotion categories (e.g., words categorized as anger), we performed the following analyses. First, we fitted the GAM model without including valence, separately for visual and auditory modalities. Second, we retrieved the visual and auditory residuals for each word and subtracted these residuals between visual and auditory modalities, yielding an auditory-visual difference score for each word. Third, we retrieved annotations of discrete emotions for as many of the included words as possible (0 indicates not associated and 1 indicates associated) from a word-emotion association database (Mohammad & Turney, 2010; Mohammad & Turney, 2013), yielding a relatively large set of words that were associated with one of five basic emotion categories: anger (669 words), disgust (552 words), fear (830 words), joy (444 words) and sadness (629 words). Finally, we tested the auditory-visual difference scores against zero for each emotion category separately.

In line with the modelling results, the auditory-visual difference was negative for words categorized as belonging to one of the negative emotions (mean difference: 0.066; t_{2679} = 3.244, p = 0.001), indicating faster responses in the auditory than the visual modality for negative words. Individual results for each negative emotion category revealed numerically the strongest auditory-visual difference for words categorized as disgust (0.152), followed by anger (0.066), sadness (0.040), and fear (0.029). When testing the auditory-visual difference for each emotion separately, only disgust was significant (t_{551} = 3.330, p < 0.001; all other emotions $p > 0.05$). However, a one-way ANOVA on the four negative emotions revealed no significant difference between the negative emotions ($F_{3,2676} = 1.68$, $p = 0.169$). Finally, as predicted, positive words (categorized as joy/happiness) showed no significant auditory-visual difference (mean = 0.051; $t_{443} = 1.041$, $p = 0.299$).

Discussion

Results from Analysis 1 showed a differential effect of valence on visual and auditory word recognition, in which the relation between valence and visual lexical decision time followed a linear function while the relation between valence and auditory lexical decision time followed an inverted-U function. By analyzing a common set of words across visual and auditory lexical decision tasks, this finding indicates that the emotional modulation of word recognition is modality-specific, thus implicating sensory-perceptual processing stages.

Though our findings support the sensory-perceptual level hypothesis, the mixed findings in prior literature on valence and visual word recognition (Estes & Adelman, 2008b; Kousta et al., 2009; Kuperman et al., 2014; Vinson et al., 2014) highlight the importance of evaluating the effects across different samples. To test whether the effects of valence on visual and auditory word recognition were similar across different samples, in Analysis 2, we merged the same set of control variables as Analysis 1, but with visual and auditory lexical decision times from five databases (i.e., ELP, ECP, BLP, MALD, and AELP) into a common dataset (4149 words).

Analysis 2: Common Words across ELP, ECP, BLP, MALD, AELP

The Effect of Valence on Visual Lexical Decision Time

To examine the relation between valence and visual word recognition across different samples, GAM was performed for visual lexical decision times from ELP, ECP and BLP (4149 words). We found that valence significantly predicted visual lexical decision times for all three databases (**Table 4**). However, the functional relationship between valence and visual lexical decision times differed across databases. Similar to Analysis 1, a linear relationship was found for ELP, here with a smaller set of words (**Figure 4A**), and a valence by frequency interaction was also found. However, for ECP and BLP, the valence effects were not linear (edfs = 2.514 and 1.992, respectively, but note that edfs do not speak directly to the specific functional form), wherein there was a slower rate of change (i.e., relatively flat) during negative valence and a faster rate of change for positive valence (**Figure 4A**). We also performed the same analyses with general frequency (Brysbaert & New, 2009) rather than written frequency was modeled, and found comparable effects of valence on visual lexical decision times (**Supplemental Figure 6A**). Given that merging across five databases limited the set size analyzed to 4149 words, in supplementary analyses, we further examined the effects on visual lexical decision times for ECP and BLP with their maximal set sizes when merged with the common set of lexical control variables (9572 words for ECP, same size as ELP; 5926 words for BLP) as Analysis 1 (**Supplemental Figure 7**).

Table 4. GAM results of major effects for visual lexical decision.

Smooth terms	edf	Ref.df	F	*p*
ELP				
Valence	0.875	9	0.770	p = 0.004
Arousal	0.002	9	F < 0.001	p = 0.434
Freq_CobW	5.595	9	85.40	p < 0.001
Valence × Arousal	1.383	81	0.031	p = 0.103

Valence × Freq_CobW	1.075	77	0.059	p = 0.020
Arousal × Freq_CobW	0.260	81	0.003	p = 0.337
ECP				
Valence	2.514	9	4.996	p < 0.001
Arousal	2.173	9	2.309	p < 0.001
Freq_CobW	6.700	9	86.72	p < 0.001
Valence × Arousal	6.328	80	0.151	p = 0.032
Valence × Freq_CobW	0.745	77	0.017	p = 0.147
Arousal × Freq_CobW	0.001	81	F < 0.001	p = 0.425
BLP				
Valence	1.992	9	1.688	p < 0.001
Arousal	1.723	9	1.123	p = 0.001
Freq_CobW	8.036	9	110.7	p < 0.001
Valence × Arousal	< 0.001	81	F < 0.001	p = 0.813
Valence × Freq_CobW	0.005	81	F < 0.001	p = 0.345
Arousal × Freq_CobW	< 0.001	81	F < 0.001	p = 0.898

Note. edf refers to effective degrees of freedom. Ref.df refers to reference degrees of freedom. The *F* is a Wald-type test statistic based on Nychka's across-the-function interpretation of Bayesian credible intervals for the spline (Wood, 2013).

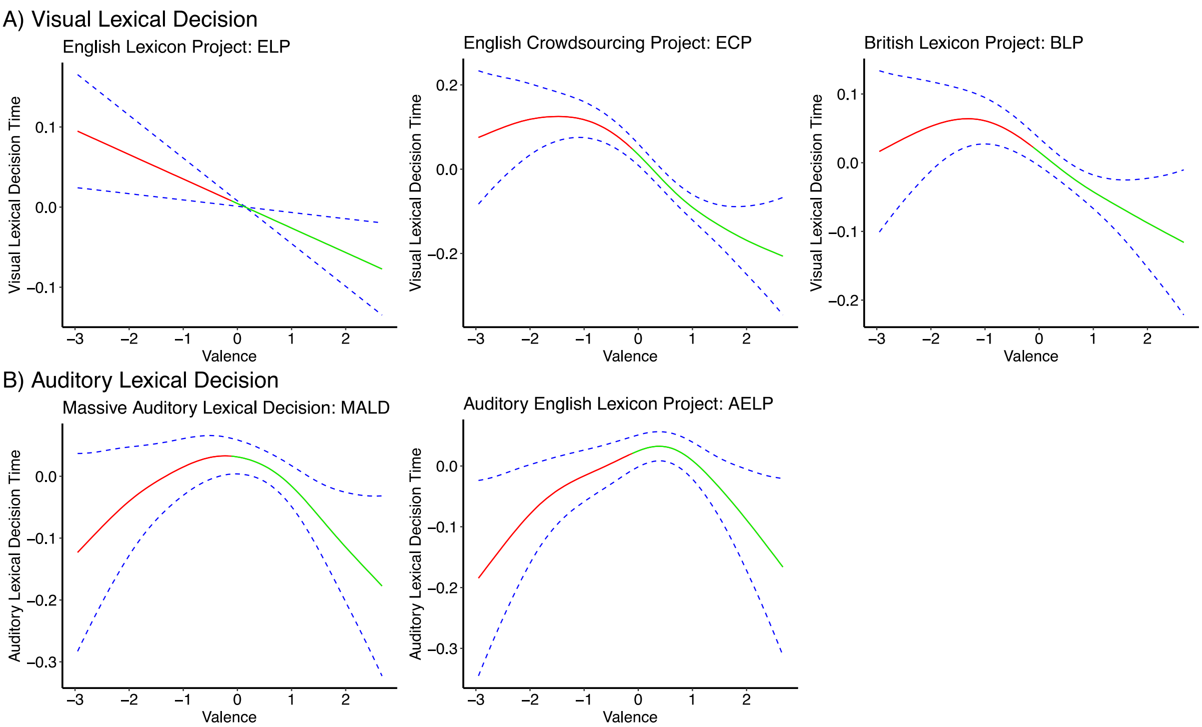

Figure 4. GAM functions describing the relationship between valence in Analysis 2 with A) visual lexical decision time from English Lexicon Project (ELP), English Crowdsourcing

Project (ECP), British Lexicon Project (BLP) and B) auditory lexical decision from Massive Auditory Lexical Decision (MALD) and Auditory English Lexicon Project (AELP). Green = Positive; red = negative valence. Dotted blue lines show 95% confidence intervals for GAM functions.

The Effect of Valence on Auditory Lexical Decision Time

To examine valence effects for auditory lexical decision tasks across different samples, GAM was performed for auditory lexical decision from MALD and AELP (4149 words). We found that valence significantly predicted auditory lexical decision times for both databases (**Table 5**). Similar to Analysis 1, the effect of valence on auditory lexical decision time followed an inverted-U function (edf = 2.070) for MALD, here with a smaller set of words, wherein positive and negative valences are nearly symmetric (**Figure 4B**). In addition, a similar inverted-U function (edf = 2.496) was identified when a different database, AELP, was used (**Figure 4B**). We also performed the same analyses when general frequency (Brysbaert & New, 2009) rather than spoken frequency was modeled, and found comparable effects of valence on auditory lexical decision time (**Supplemental Figure 6B**). Given that merging across five databases limited the set size analyzed to 4149 words, we further performed supplementary analyses to examine the effects on auditory lexical decision times for AELP with its maximal set size (6236 words) when merged with a common set of predictors (**Supplemental Figure 7**).

Table 5. GAM results of major effects for auditory lexical decision.

Smooth terms	edf	Ref.df	F	*p*
MALD				
Valence	2.070	9	0.829	p = 0.014
Arousal	0.909	9	1.076	p < 0.001
Freq_CobS	1.961	9	3.197	p < 0.001
Valence × Arousal	0.005	81	F < 0.001	p = 0.796
Valence × Freq_CobS	3.130	81	0.083	p = 0.040
Arousal × Freq_CobS	1.739	80	0.031	p = 0.171
AELP				
Valence	2.496	9	1.011	p = 0.008
Arousal	0.958	9	2.002	p < 0.001
Freq_CobS	3.355	9	29.57	p < 0.001
Valence × Arousal	1.642	80	0.035	p = 0.107
Valence × Freq_CobS	1.244	31	0.098	p = 0.065
Arousal × Freq_CobS	1.308	80	0.037	p = 0.062

Note. edf refers to effective degrees of freedom. Ref.df refers to reference degrees of freedom. The *F* is a Wald-type test statistic based on Nychka's across-the-function interpretation of Bayesian credible intervals for the spline (Wood, 2013).

Comparing the Effects of Valence on Visual and Auditory Lexical Decision Times

To further statistically examine whether the valence effects were asymmetric for visual but symmetric for auditory, *Valence* was changed from a continuous variable to a binary variable (positive, negative) using median split. We used the dataset with a set of common words (4149 words) with lexical decision times of the five databases (i.e., ELP, ECP, BLP, MALD, and AELP), which allowed us to compare the effects with equal sample size across databases. Note that general frequency rather than written or spoken frequency was used in this analysis to have a common frequency variable across visual and auditory. We did not include the *Modality* factor for this analysis to test the interaction between *Modality* and *Valence* because there were three visual databases and two auditory databases.

Analyses were performed for each of the five databases depending on visual or auditory lexical decision time, with Equation 5 for the visual and Equation 6 for the auditory modality (in section Analysis 1, Equations 5-6).

The valence effect was significant for ELP ($t = -2.168$, $p = 0.030$), ECP ($t = -7.755$, $p < 0.001$), and BLP ($t = -4.544$, $p < 0.001$) but not for MALD ($t = 0.930$, $p = 0.353$) or AELP ($t = -0.266$, $p = 0.790$). These results were comparable to those of Analysis 1, suggesting that the effects of valence on visual lexical decision were asymmetric (i.e., negativity slow-down effects) while the effects of valence on auditory lexical decision were symmetric. We did not examine the effects for distinct emotion categories for Analysis 2 because there were three visual databases and two auditory databases.

Discussion

Results from Analysis 2 revealed the effects of valence on visual and auditory lexical decision times across five databases: ELP, ECP, BLP, MALD, and AELP on a common set of words. We found that the effect of valence on visual lexical decision time was linear for ELP, and the effect of valence on auditory lexical decision time was inverted-U for MALD, replicating Analysis 1 with a smaller set of words. The effects of valence on visual lexical decision time for ECP and BLP were similar to each other, with a slower rate of change during negative valence and a faster rate of change during positive valence. Although there were differences between ELP, ECP, and BLP databases, visual lexical decision times in all three databases were slower for negative than positive words. By contrast, the effects of valence on auditory lexical decision time followed an inverted-U function, for both MALD and AELP databases. Unlike the asymmetric effects observed for visual lexical decision time, valence had a symmetric effect on auditory lexical decision time, with faster responses for both positive and negative words.

The findings of Analysis 2 were in line with those of Analysis 1. However, it also demonstrated differences across databases, even for the same set of words. While some of these differences may be meaningful, they could also reflect differences that are not of interest. Therefore, in Analysis 3 we averaged visual lexical decision times across ELP, ECP and BLP, and averaged auditory lexical decision times across MALD and AELP.

Analysis 3: Averaged Across Visual or Auditory Lexical Decision Times

The Effect of Valence on Visual Lexical Decision Time

Valence significantly predicted visual lexical decision time (**Table 6**). Similar to the findings for ECP and BLP in Analysis 2, the valence effect on average visual lexical decision time was not linear (edf = 2.243), with a slower rate of change during negative valence and a faster rate of change for positive valence. Negative valence generally slowed down visual lexical decision time compared to positive valence (**Figure 5A**). Functional relationships of other significant predictors with visual lexical decision time are shown in **Supplemental Figure 8A**. We also performed the same analyses when general frequency (Brysbaert &

New, 2009) rather than written frequency was modeled, and found comparable effects of valence on visual lexical decision time (**Supplemental Figure 9**).

Table 6. GAM results of major effects for visual and auditory lexical decision.

Smooth terms	edf	Ref.df	F	*p*
Visual				
Valence	2.243	9	3.840	$p < 0.001$
Arousal	2.158	9	1.875	$p < 0.001$
Freq_CobW	7.707	9	153.5	$p < 0.001$
Valence × Arousal	0.002	81	$F < 0.001$	$p = 0.441$
Valence × Freq_CobW	0.776	74	0.027	$p = 0.087$
Arousal × Freq_CobW	0.001	81	$F < 0.001$	$p = 0.741$
Auditory				
Valence	2.815	9	1.649	$p < 0.001$
Arousal	0.958	9	2.413	$p < 0.001$
Freq_CobS	2.903	9	19.31	$p < 0.001$
Valence × Arousal	0.488	80	0.008	$p = 0.249$
Valence × Freq_CobS	2.649	81	0.087	$p = 0.021$
Arousal × Freq_CobS	0.006	81	$F < 0.001$	$p = 0.453$

Note. edf refers to effective degrees of freedom. Ref.df refers to reference degrees of freedom. The null distributions of F were non-standard.

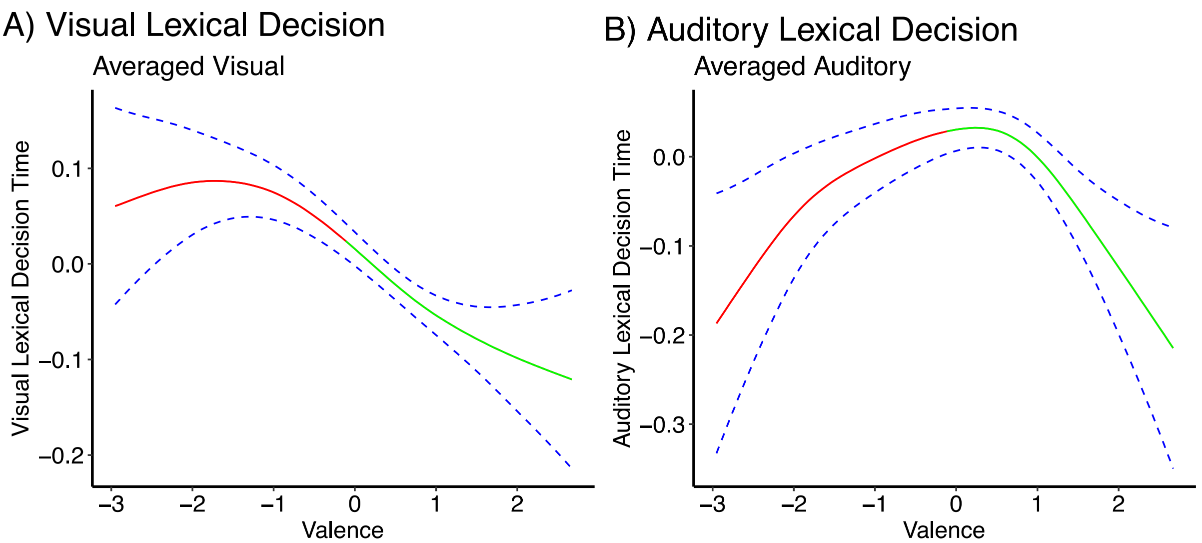

Figure 5. GAM functions describing the relationship between valence in Analysis 3 with A) averaged visual lexical decision time and B) averaged auditory lexical decision time. Green = Positive valence; red = negative valence. Dotted blue lines show 95% confidence intervals for GAM functions.

The Effect of Valence on Auditory Lexical Decision Time

Next, we examined the effect of valence on auditory lexical decision time for the same set of words, now averaging data across MALD and AELP. We found that valence significantly predicted auditory lexical decision time (**Table 6**). Similar to Analyses 1 and 2, the effect of valence on auditory lexical decision time was not linear (edf = 2.815). In contrast to the asymmetry found for visual lexical decision time (negative>positive), the effect of valence on auditory lexical decision time was symmetric (**Figure 5B**). Functional relationships of other significant predictors with auditory lexical decision time are shown in **Supplemental Figure 8B**. We also performed the same analyses modeling general frequency (Brysbaert & New, 2009) rather than spoken frequency, and found comparable effects of valence on auditory lexical decision time (**Supplemental Figure 9**).

Comparing the Effects of Valence on Visual and Auditory Lexical Decision Times

To statistically contrast the influence of valence on visual versus auditory lexical decision, we performed two analyses with *Modality* (visual, auditory) included as factor, as in Analysis 1. We used the dataset with a set of common words (4149 words) with visual lexical decision time averaged across ELP, ECP and BLP, and auditory lexical decision time averaged across MALD and AELP. The lexical decision times were $\log_{10}$-transformed, and standardized to have a mean of 0 and standard deviation of 1 before being averaged (see Method section).

First, to evaluate whether the functional forms are different for the visual versus the auditory modality, we examined if there was an interaction between *Modality* and *Valence* by including both variables into one overall model. Note that general frequency rather than written or spoken frequency was used in this analysis to have a common frequency variable across visual and auditory lexical decision times. The model structure for this analysis was as Equation 3.

A significant interaction between valence and modality was found ($F_{edf=2.98,\ Ref.df=9}$ = 1.511, $p = 0.002$), indicating a differential influence of valence on visual and auditory lexical decision times.

Second, to further statistically examine whether the valence effects were asymmetric for visual but symmetric for auditory, *Valence* was changed from a continuous variable to a binary variable (positive, negative) using median split as in Analyses 1 and 2.

The model structure for this analysis was as Equation 4. Follow-up analyses were then performed for visual and auditory lexical decision times with Equations 5 and 6, respectively (in section Analysis 1, Equations 4-6).

A significant interaction between valence and modality was found ($t = 6.908$, $p < 0.001$), indicating a differential influence of valence on visual and auditory lexical decision times. The valence effect was significant for visual ($t = -6.344$, $p < 0.001$) but not for auditory ($t = -0.767$, $p = 0.443$). These results were comparable to those of Analyses 1 and 2, suggesting that the effects of valence on visual lexical decision were asymmetric (i.e., negativity slow-down effects) while the effects of valence on auditory lexical decision were symmetric.

Effects for Distinct Emotion Categories

Similar to Analyses 1 and 2, the above analyses showed that the effects of positive valence were similar across modalities while the effects of negative valence were different. As in Analysis 1, we again tested whether the negative valence difference was driven by specific negative emotion categories, following the procedure described in Analysis 1. The number of words that were associated with one of five basic emotion categories were: anger (300 words), disgust (237 words), fear (359 words), joy (219 words) and sadness (271 words).

In line with the modelling results, the auditory-visual difference was negative for words categorized as belonging to one of the negative emotions (mean difference: 0.0494; t_{1166} = 2.132, p = 0.033), indicating faster responses in the auditory than the visual modality for negative words. Individual results for each negative emotion category revealed numerically the strongest auditory-visual difference for words categorized as disgust (0.101), followed by sadness (0.0438), fear (0.0418), and anger (0.0226). Similar to Analysis 1, when testing the auditory-visual difference for each emotion separately, only disgust was significant (t_{236} = 1.990, p = 0.047; all other emotions p > 0.05). However, again, a one-way ANOVA on the four negative emotions revealed no significant difference between the negative emotions ($F_{3,1163}$ = 0.467, p = 0.705). Finally, as predicted, positive words (categorized as joy/happiness) showed no significant auditory-visual difference (mean = -0.0125; t_{218} = -0.237, p = 0.813).

Discussion

Results from Analysis 3 showed a differential effect of valence on visual and auditory word recognition when averaging across visual or auditory lexical decision times from multiple databases. In line with Analysis 2, the valence effect for visual word recognition again showed slower responses to negative than positive words (referred to as the negativity bias in previous literature), whereas the valence effect for auditory word recognition showed a symmetric effect.

General Discussion

The main goal of the present study was to determine the effects of valence on visual and auditory word recognition, as measured by a lexical decision task. Previous work primarily investigated the effects of valence on visual word recognition, such that it was unknown how valence influences auditory word recognition and whether the effects of valence are different or similar across modalities. Here, by examining large samples of words from multiple visual and auditory databases, we found that valence differentially influenced visual and auditory word recognition: while the effects of valence on visual lexical decision time were asymmetric, the effects of valence on auditory lexical decision time were symmetric, following an inverted-U function. The modality specificity of valence effects was consistent across multiple analyses and multiple large databases. However, while the overall functional relationship between valence and lexical decision times differed across modalities, the effects of positive valence were consistent across modalities: for both modalities we found that the more positive the valence of the word, the faster the lexical decision time. By contrast, the effect of negative valence was strikingly different across modalities, with negative valence facilitating lexical decision only in the auditory modality.

We interpret these findings as evidence that the effects of valence on word recognition partly occur at the sensory-perceptual level, considering that post-perceptual processes (e.g., semantic processes) are shared between modalities (Caramazza & Hillis, 1990; McClelland & Rumelhart, 1981). The effect of valence on word recognition may be a specific instantiation of the effect of reinforcers on sensory processing (Anderson, 2019), with emotions being closely related to rewards and punishments (Rolls, 1990). Specifically, previous work has shown that visual stimuli (e.g., objects) that become associated with positive outcome (e.g., monetary gain) are more rapidly recognized (Anderson et al., 2011; Becker et al., 2020; Hickey et al., 2015) and are more strongly represented in visual cortex than neutral stimuli (Hickey & Peelen, 2015, 2017; Serences & Saproo, 2010). Similarly, the positive meaning of a word (e.g. "money") may act as a positive reinforcer, facilitating the sensory processing of the word. This is consistent with literature showing relatively early (as early as before or around 200 ms) ERP effects of valence on word processing (Begleiter & Platz, 1969; Bernat et al., 2001; Fritsch & Kuchinke, 2013; Herbert et al., 2008; Kaltwasser et al., 2013; Kanske & Kotz, 2007; Kissler & Herbert, 2013; Landis, 2006; Ortigue et al., 2004; Schacht et al., 2012; Scott et al., 2009; Stolarova et al., 2006). There is some evidence that positive reinforcement similarly facilitates auditory processing (Anderson, 2016; Asutay & Västfjäll, 2016), in line with our results of similar effects of positive valence on visual and auditory word recognition.

Conversely, the negative meaning of a word (e.g., "poverty") may act as a negative reinforcer. While the effect of negative reinforcement (e.g., monetary loss) on sensory processing has not been studied as much as the effect of positive reinforcement, there is evidence that loss-associated visual stimuli are detected less accurately than neutral stimuli (Barbaro et al., 2017; But see Wentura et al., 2014), and that their representations in visual cortex are suppressed relative to neutral stimuli (Barbaro et al., 2017). This effect may be similar to the suppressive effects observed for disgust-related stimuli (Chapman & Anderson, 2012; Krusemark & Li, 2011) but different from the facilitatory effect observed for threat-related stimuli (Krusemark & Li, 2011). A suppressive effect of negative association on visual recognition is in line with the current findings, particularly with the results of the ELP database (Analysis 1), showing slower reaction times for negative words. Intriguingly, we found that negative association facilitated word recognition in the auditory modality, with faster responses to negative words. This is in line with the rapid processing of disgust vocalizations (Sauter & Eimer, 2010). However, we are not aware of studies experimentally associating auditory stimuli with negative outcome (e.g., monetary loss). The current results suggest that the processing of such stimuli may be facilitated in the auditory domain, a prediction that can be tested in future work.

The finding of faster recognition of negative words in the auditory (but not visual) modality can also be interpreted in the context of the proposal that the auditory system functions as an "early warning system" (Murphy et al., 2013): Unlike visual processing, auditory processing is not restricted to fixated locations in space, making the auditory system highly suitable for detecting potentially threatening objects or events from all directions, even in the dark. Accordingly, the rapid processing of negative auditory words observed here may reflect this function. On this account, we may expect a particularly strong auditory-visual difference for words signaling danger (i.e., words categorized as anger or fear). The analysis of distinct emotion categories did not directly support this hypothesis, however, showing no significant difference between the four negative emotions (sadness, disgust, anger, fear), and showing numerically the strongest auditory-visual difference for words categorized as disgust.

Another inherent difference between auditory and visual word recognition is that auditory information unfolds across time while visual information is instantly available. It is possible, therefore, that the differential effect of valence on auditory and visual word recognition revealed here is due to this temporal difference. For example, the longer processing time for auditory words may give more or deeper access to semantic representations. However, it is unclear why this would differentially affect positive and negative words. To test the influence of the temporal component, future studies could reveal visual words gradually, at the same speed as auditory words, and measure the effect of valence on this temporally-matched visual lexical decision task.

Within the visual modality, our study adds to previous studies investigating the effect of valence on word recognition. We consistently found an asymmetric effect of valence, in which negative valence generally slowed down reaction time compared to positive valence (Gao et al., 2020; Pratto & John, 1991). This is consistent with prior literature that found the relation between valence and word recognition to be either a step (Estes & Adelman, 2008a; 2008b; Figure 1A) or a monotonic (Kuperman et al., 2014; Figure 1C) function, while contradicting studies that found a symmetric function (Kousta et al., 2009; Vinson et al., 2014; Figure 1B). These differences can be due to the size of the included dataset, the databases analyzed, or lack of control variables, such as valence by frequency interactions. For example, the majority of previous studies included about 1000 words. A recent study that included more than 10.000 words revealed a monotonic near-linear effect of valence (Kuperman et al., 2014), similar to results of Analysis 1 that used the same database. One study highlighted the issue of statistical power on how valence influences lexical decision times, especially in factorial experimental designs when discretizing continuous valence into a discrete variable (Kuperman, 2015). Intriguingly, in further analyses, we found that the effect of valence on visual word recognition differed across databases, even when including the same words and control variables. While a linear function was found for ELP, a nonlinear function (but monotonic negative) was found for ECP and BLP, in which negative valence generally slowed down reaction time compared to positive valence but there was a slower rate of change during negative valence compared to positive valence. These findings may explain differences between previous studies that used one of the databases. For example, using over 1,000 words from British English Lexicon Project, Vinson et al. (2014) found a symmetric valence effect, which is different from findings of other studies using English Lexicon Project with comparable dataset sizes (Estes & Adelman, 2008b; Larsen et al., 2008). Using a bootstrapping approach, Kuperman (2015) also showed that the probabilities of finding a linear or inverted-U function for the effect of valence on visual lexical decision time was affected by whether stimuli were selected from ELP or BLP. It is still unclear why reaction times differ systematically across databases. Nevertheless, such differences highlight the importance of evaluating the effects of valence on word recognition using multiple databases. The average of the three databases may provide the best estimate (Figure 5).

The current study is the first large-scale study investigating the effects of valence on auditory lexical decision. Different from the visual modality, we found that the effects of valence on auditory lexical decision were symmetric: both negative and positive valence facilitated word recognition. In the visual modality, such symmetric effects have been interpreted as evidence for the motivated attention theory (Kousta et al., 2009; Vinson et al., 2014). According to this theory (Lang et al., 1990), there are two motivation-emotion systems: appetitive and defensive. Positive valence activates the appetitive motivational system, preparing the organism for approach behaviors; whereas negative valence activates the defensive motivational system, leading to avoidance behaviors. Positive and negative valences are equally important for preparing humans for successful behaviors, and therefore, this theory predicts that the relation between valence and word recognition follows a symmetric function. The current auditory results can be interpreted as supporting the motivated attention theory, though it is unclear why visual word recognition would not follow a similar pattern of results.

In conclusion, our findings demonstrate a differential effect of valence on visual and auditory word recognition. The effect of valence on visual word recognition was asymmetric, in which positive valence facilitated word recognition. By contrast, the effect of valence on auditory word recognition was symmetric, with both positive and negative valence facilitating word recognition relative to neutral words. We interpret these findings as indicating that valence influences word recognition partly at the sensory-perceptual stage, similar to effects of positive and negative reinforcers on sensory processing more generally (e.g., object recognition). Altogether, our findings extend the investigation of valence effects on word recognition from a single visual modality to both visual and auditory modalities, opening up new directions for future research.

References

Anderson, B. A. (2016). Value-driven attentional capture in the auditory domain. *Attention, Perception, & Psychophysics, 78*(1), 242-250.

Anderson, B. A. (2019). Neurobiology of value-driven attention. *Current opinion in psychology, 29,* 27-33.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367-10371.

Asutay, E., & Västfjäll, D. (2016). Auditory attentional selection is biased by reward cues. *Scientific Reports, 6*(1), 1-6.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1996). The CELEX lexical database (cd-rom).

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*(3), 340.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*(2), 283.

Balota, D. A., Yap, M. J., & Cortese, M. J. (2006). Visual word recognition: The journey from features to meaning (a travel update). In *Handbook of psycholinguistics* (pp. 285-375). Elsevier.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445-459.

Barbaro, L., Peelen, M. V., & Hickey, C. (2017). Valence, not utility, underlies reward-driven prioritization in human vision. *Journal of Neuroscience, 37*(43), 10438-10450.

Becker, M. W., Hemsteger, S. H., Chantland, E., & Liu, T. (2020). Value-based attention capture: Differential effects of loss and gain contingencies. *Journal of Vision, 20*(5), 4-4.

Begleiter, H., & Platz, A. (1969). CORTICAL EVOKED POTENTIALS TO SEMANTIC STIMULI. *Psychophysiology*, *6*(1), 91-100. https://doi.org/https://doi.org/10.1111/j.1469-8986.1969.tb02887.x

Bernat, E., Bunce, S., & Shevrin, H. (2001, 2001/08/01/). Event-related brain potentials differentiate positive and negative mood adjectives during both supraliminal and subliminal visual processing. *International Journal of Psychophysiology*, *42*(1), 11-34. https://doi.org/https://doi.org/10.1016/S0167-8760(01)00133-7

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977-990.

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods, 44*(4), 991-997.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904-911.

Caramazza, A., & Hillis, A. E. (1990). Where do semantic errors come from? *Cortex, 26*(1), 95-122.

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*(4), 857.

Chapman, H. A., & Anderson, A. K. (2012). Understanding disgust. In *The year in cognitive neuroscience.* (pp. 62-76). Blackwell Publishing.

Deniz, F., Nunez-Elizalde, A. O., Huth, A. G., & Gallant, J. L. (2019). The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. *Journal of Neuroscience, 39*(39), 7722-7736.

Estes, Z., & Adelman, J. S. (2008a). Automatic vigilance for negative words in lexical decision and naming: Comment on Larsen, Mercer, and Balota (2006).

Estes, Z., & Adelman, J. S. (2008b). Automatic vigilance for negative words is categorical and general.

Ferrand, L., Méot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., Dufau, S., Mathôt, S., & Grainger, J. (2018). MEGALEX: A megastudy of visual and auditory word recognition. *Behavior Research Methods, 50*(3), 1285-1307.

Fritsch, N., & Kuchinke, L. (2013). Acquired affective associations induce emotion effects in word recognition: an ERP study. *Brain and Language, 124*(1), 75-83.

Gao, C., Wedell, D. H., & Shinkareva, S. V. (2020, Nov 19). Crossmodal negativity bias in semantic processing. *Emotion*. https://doi.org/10.1037/emo0000918

Gerhand, S., & Barry, C. (1999). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory & Cognition, 27*(4), 592-602.

Goh, W. D., Yap, M. J., & Chee, Q. W. (2020). The Auditory English Lexicon Project: A multi-talker, multi-region psycholinguistic database of 10,170 spoken words and nonwords. *Behavior Research Methods*, 1-30.

Goh, W. D., Yap, M. J., Lau, M. C., Ng, M. M., & Tan, L.-C. (2016). Semantic richness effects in spoken word recognition: A lexical decision and semantic categorization megastudy. *Frontiers in Psychology, 7*, 976.

Goldinger, S. D. (1996). Auditory lexical decision. *Language and Cognitive Processes, 11*(6), 559-568.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language, 24*(6), 631-645.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29*(2), 228-244.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychological review, 103*(3), 518.

Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized additive models* (Vol. 43). CRC press.

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology, 45*(3), 487-498.

Hickey, C., Kaiser, D., & Peelen, M. V. (2015). Reward guides attention to object categories in real-world scenes. *Journal of Experimental Psychology: General, 144*(2), 264.

Hickey, C., & Peelen, M. V. (2015). Neural mechanisms of incentive salience in naturalistic human vision. *Neuron, 85*(3), 512-518.

Hickey, C., & Peelen, M. V. (2017). Reward selectively modulates the lingering neural representation of recently attended objects in natural scenes. *Journal of Neuroscience, 37*(31), 7297-7304.

Hudson, P. T., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory and Language, 24*(1), 46-58.

Jacobs, A. M., & Grainger, J. (1994). Models of visual word recognition: sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1311.

Jessen, F., Heun, R., Erb, M., Granath, D.-O., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The concreteness effect: Evidence for dual coding and context availability. *Brain and Language, 74*(1), 103-112.

Kahan, T. A., & Hely, C. D. (2008). The role of valence and frequency in the emotional Stroop task. *Psychonomic bulletin & review, 15*(5), 956-960.

Kaltwasser, L., Ries, S., Sommer, W., Knight, R., & Willems, R. M. (2013). Independence of valence and reward in emotional word processing: electrophysiological evidence. *Frontiers in Psychology, 4*, 168.

Kanske, P., & Kotz, S. A. (2007, 2007/05/07/). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research, 1148*, 138-148. https://doi.org/https://doi.org/10.1016/j.brainres.2007.02.044

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44*(1), 287-304.

Kissler, J., & Herbert, C. (2013). Emotion, Etmnooi, or Emitoon?–Faster lexical access to emotional than to neutral words during reading. *Biological Psychology, 92*(3), 464-479.

Kissler, J., & Koessler, S. (2011). Emotionally positive stimuli facilitate lexical decisions—An ERP study. *Biological Psychology, 86*(3), 254-264.

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*(3), 473-481.

Krusemark, E. A., & Li, W. (2011). Do all threats work the same way? Divergent effects of fear and disgust on sensory perception and attention. *Journal of Neuroscience, 31*(9), 3429-3434.

Kuperman, V. (2015). Virtual experiments in megastudies: a case study of language and emotion. *Quarterly Journal of Experimental Psychology, 68*(8), 1693-1710.

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General, 143*(3), 1065.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978-990.

Landis, T. (2006, 2006/01/01/). Emotional Words: What's so Different from Just Words? *Cortex, 42*(6), 823-830. https://doi.org/https://doi.org/10.1016/S0010-9452(08)70424-6

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological review, 97*(3), 377.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006, Feb). Lexical characteristics of words used in emotional Stroop experiments. *Emotion, 6*(1), 62-72. https://doi.org/10.1037/1528-3542.6.1.62

Larsen, R. J., Mercer, K. A., Balota, D. A., & Strube, M. J. (2008). Not all negative words slow down lexical decision and naming speed: Importance of word arousal.

Mandera, P., Keuleers, E., & Brysbaert, M. (2019). Recognition times for 62 thousand English words: Data from the English Crowdsourcing Project. *Behavior Research Methods*, 1-20.

Marra, G., & Wood, S. N. (2011). Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis, 55*(7), 2372-2387.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review, 88*(5), 375.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of experimental psychology, 90*(2), 227.

Mohammad, S., & Turney, P. (2010). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text,

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational intelligence, 29*(3), 436-465.

Murphy, S., Fraenkel, N., & Dalton, P. (2013). Perceptual load does not modulate auditory distractor processing. *Cognition, 129*(2), 345-355.

Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *Journal of classification, 31*(3), 274-295.

Ortigue, S., Michel, C. M., Murray, M. M., Mohr, C., Carbonnel, S., & Landis, T. (2004, 2004/04/01/). Electrical neuroimaging reveals early generator modulation to emotional words. *Neuroimage, 21*(4), 1242-1251. https://doi.org/https://doi.org/10.1016/j.neuroimage.2003.11.007

Pratto, F., & John, O. P. (1991). Automatic vigilance: the attention-grabbing power of negative social information. *Journal of personality and social psychology, 61*(3), 380.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological review, 111*(1), 159.

Rogowsky, B. A., Calhoun, B. M., & Tallal, P. (2016). Does modality matter? The effects of reading, listening, and dual modality on comprehension. *SAGE Open, 6*(3), 2158244016669550.

Rolls, E. T. (1990). A theory of emotion, and its application to understanding the neural basis of emotion. *Cognition & Emotion, 4*(3), 161-190.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological review, 89*(1), 60.

Sauter, D. A., & Eimer, M. (2010). Rapid detection of emotion from human vocalizations. *Journal of Cognitive Neuroscience, 22*(3), 474-481.

Schacht, A., Adler, N., Chen, P., Guo, T., & Sommer, W. (2012). Association with positive outcome induces early effects in event-related brain potentials. *Biological Psychology, 89*(1), 130-136.

Scott, G. G., O'Donnell, P. J., Leuthold, H., & Sereno, S. C. (2009). Early emotion word processing: Evidence from event-related potentials. *Biological Psychology, 80*(1), 95-104.

Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension—an fMRI study. *Cerebral Cortex, 22*(7), 1662-1670.

Serences, J. T., & Saproo, S. (2010). Population response profiles in early visual cortex are biased in favor of more valuable stimuli. *Journal of neurophysiology, 104*(1), 76-87.

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: an eye-movement study. *Emotion, 13*(6), 1107.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science, 8*(6), 411-416.

Stolarova, M., Keil, A., & Moratti, S. (2006). Modulation of the C1 visual event-related component by conditioned stimuli: evidence for sensory plasticity in early affective perception. *Cerebral Cortex, 16*(6), 876-887.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124*(2), 107.

Tucker, B. V., Brenner, D., Danielson, D. K., Kelley, M. C., Nenadić, F., & Sims, M. (2019). The massive auditory lexical decision (MALD) database. *Behavior Research Methods, 51*(3), 1187-1204.

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion, 28*(4), 737-746.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191-1207.

Wentura, D., Müller, P., & Rothermund, K. (2014). Attentional capture by evaluative stimuli: Gain- and loss-connoting colors boost the additional-singleton effect. *Psychonomic bulletin & review, 21*(3), 701-707.

Wood, S., & Wood, M. S. (2015). Package ‘mgcv’. *R package version, 1*, 29.

Wood, S. N. (2013). On p-values for smooth components of an extended generalized additive model. *Biometrika, 100*(1), 221-228.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC press.

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language, 72*, 37-48.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language, 60*(4), 502-529.

Yap, M. J., & Seow, C. S. (2014). The influence of emotion on lexical processing: Insights from RT distributional analysis. *Psychonomic bulletin & review, 21*(2), 526-533.

Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language, 48*(4), 779-793.