Evaluation of lexical and semantic features for English emotion words

Francesca Citron, Kathrin A.M. Klingebiel, Brendan S. Weekes and Evelyn C. Ferstl

University of Sussex, Department of Psychology, Falmer, Brighton, UK - F.Citron@sussex.ac.uk

Introduction

Lexical processing of written words is the base for constructing meaning from text, as words are the primary meaningful elements provided to the reader. Words are characterized by many different lexical-semantic features which affect verbal processing and hence behavioral performance as reaction times, eye-movements, fixation times and brain activity, as shown by ERP and neuroimaging. We distinguish here lexical features as word frequency, familiarity, age of acquisition, from semantic ones as imageability and concreteness ("cold features") and from affective ones as emotional valence and arousal ("warm features"). In particular, emotional valence (positive, negative) has a strong early impact on cognitive processing (Posner et al., 2007; Sreenivasan et al., 2009) and tends to modulate emotional information are processed in a qualitatively different way when words containing emotional content are presented (Ferland et al., 2007). These findings provide an impetus to investigate the role of emotional features in the comprehension of words, which is important for the understanding of complex stimuli which are made of many different features of words, but also features related to emotional connotations.

Aim

The aim of the present study was to generate a corpus of English words of different emotional valence, evaluated for various lexical and semantic features. This corpus will be useful for experiments employing ERPs and eye-tracking techniques, to allow a better control of the effects of these features. Furthermore, the corpus could allow a well-balanced selection of words to be used for discourse processing research, so that the effect of text features. We distinguish here lexical features as word frequency, familiarity, age of acquisition, from semantic ones as imageability and concreteness ("cold features") and from affective ones as emotional valence and arousal ("warm features"). Furthermore, the corpus could allow a well-balanced selection of words to be used for discourse processing research, so that the effect of text features. The corpus will be useful for experiments employing ERPs and neuroimaging. We distinguish here lexical features as word frequency, familiarity, age of acquisition, from semantic ones as imageability and concreteness ("cold features") and from affective ones as emotional valence and arousal ("warm features"). Further, the corpus could allow a well-balanced selection of words to be used for discourse processing research, so that the effect of text features. The corpus will be useful for experiments employing ERPs and neuroimaging. We distinguish here lexical features as word frequency, familiarity, age of acquisition, from semantic ones as imageability and concreteness ("cold features") and from affective ones as emotional valence and arousal ("warm features").

Results

Valence categories (positive, negative, neutral) were created by grouping the valence scores. A dummy variable called "Absoloute Valence" was created to obtain low and high emotionality ratings independent of valence.

Reliability analysis

The MRC Psycholinguistic database contains familiarity and imageability ratings for 181 out of our 300 words and AoA ratings for 72 words. Sparseness correlations between the MRC ratings and the present ones were highly significant (Fam: rho=.786, p<.001; Imag-rho=.608, p<.001; AoA-rho=.524, p<.001), showing high reliability of our ratings with previous ratings. The present study was to generate a corpus of English words of different emotional valence, evaluated for various lexical and semantic features. This corpus will be useful for experiments employing ERPs and eye-tracking techniques, to allow a better control of the effects of these features. Furthermore, the corpus could allow a well-balanced selection of words to be used for discourse processing research, so that the effect of text features. The corpus is an important tool for the understanding of complex stimuli which are made of many different features of words, but also features related to emotional connotations.

Discussion

Familiarity, age of acquisition, frequency of use Familiarity, AoA and frequency of use were all highly correlated to one another, suggesting that words which are also frequently used and were acquired early. These results support previous findings (Morrison et al., 1997; Stadthagen-Gonzales et al., 2006).

Valence and Arousal

Absolute valence and arousal were highly correlated to one another, suggesting that words which are also frequently used and were acquired early. These results support previous findings (Morrison et al., 1997; Stadthagen-Gonzales et al., 2006).

Emotionality ratings were usually higher for positive valence and arousal, but low for negative valence and arousal. A negative correlation between emotional valence and arousal suggests that high-valenced words are naturally more arousing than highly positive ones, as very intense positive feelings like "excitement" would transform into negative ones (just as saw us earlier for different findings).

Emotionality and frequency

This correlation suggests that positive words are more familiar. This could be due to the fact that positive words tend to be earlier acquired than negative ones (as shown by a weak correlation to AoA). Alternatively, there might be a misfire bias for negative words: Who wants to admit that they are very familiar with "miserable"?

Imageability and Concreteness

Imageability and concreteness are highly correlated to one another and show a similar control over emotional valence. Arousal and Imageability

This correlation suggests that high-valenced words are slightly more arousing than test imageable words. Only imageability but not valence correlated with AoA, which again is closely related to human picturing activity while reading words, in particular the context of emotional features (Parisio et al., 1998).

Imageability and age of acquisition

Imageability and concreteness are correlated, suggesting that highly imageable words are acquired earlier than less imageable words (Morrison et al., 1997; Bird et al., 2001).