To appear in Cognition. This preprint may contain errors not present in the published version.

Evolution of emotion semantics

Aotao Xu[a,*], Jennifer E. Stellar[b], Yang Xu[a,c]

[a] *Department of Computer Science, University of Toronto*
[b] *Department of Psychology, University of Toronto*
[c] *Cognitive Science Program, University of Toronto*

Abstract

Humans possess the unique ability to communicate emotions through language. Although concepts like anger or awe are abstract, there is a shared consensus about what these English emotion words mean. This consensus may give the impression that their meaning is static, but we propose this is not the case. We cannot travel back to earlier periods to study emotion concepts directly, but we can examine text corpora, which have partially preserved the meaning of emotion words. Using natural language processing of historical text, we found evidence for semantic change in emotion words over the past century and that varying rates of change were predicted in part by an emotion concept's prototypicality—how representative it is of the broader category of "emotion". Prototypicality negatively correlated with historical rates of emotion semantic change obtained from text-based word embeddings, beyond more established variables including usage frequency in English and a second comparison language, French. This effect for prototypicality did not consistently extend to the semantic category of birds, suggesting its relevance for predicting semantic change may be category-dependent. Our results suggest emotion semantics are evolving over time, with prototypical emotion words remaining semantically stable, while other emotion words evolve more freely.

Keywords: emotion; semantic change; semantic stability; prototype theory; word embedding

*Corresponding author
Email address: `a26xu@cs.toronto.edu` (Aotao Xu)

1. Introduction

Much like emotion concepts vary in their meaning across cultures [1, 2, 3], it is possible emotion words can take on different meanings over time.[1] For instance, the English word *awe* in the 18th century may not represent the same feeling it does today, after a century of evolving perspectives on power and beauty [4]. Although we cannot travel to earlier historical periods to study emotion concepts directly, we do have access to text corpora which have partially preserved the meaning of emotion words. These words do not reflect the entirety of an emotion concept, which includes expressive, experiential, and physiological components, but they do offer insight into its shared meaning within a society. Here we use computational linguistic analyses to investigate the evolution of emotion semantics.

If the meaning of different emotion words like *awe* or *joy* are changing over time, are they changing at the same rate or are there features of an emotion word that might predict its rate of semantic change? We propose that an emotion's conceptual prototypicality is one such feature. Prototypicality is a graded measure of the goodness of a concept's membership in a semantic category [5, 6]. In the case of emotions, *joy* is considered a more prototypical concept than *optimism*. Prototypical emotion concepts may have clearer biological and cultural functions and more distinct features than less prototypical ones. For instance, prototypical concepts like *fear* and *disgust* are particularly suited to solving evolutionary challenges or taking advantage of opportunities that faced early humans [7], and they may have particularly strong social or cultural scripts [8, 5, 9]. These emotion concepts are often more clearly marked by distinctive expressions, experience, and patterns of activation in the body [10], and prototypical members may even help define the meaning of their less proto-

[1] In our study, we use the terms "emotion concept" and "emotion word" interchangeably to refer to emotions that are lexicalized in natural language.

typical counterparts [11] (see *Supplementary Information* for further evidence). We hypothesize that these well-defined functions and features of prototypical emotion concepts could promote semantic stability. As a result, the meaning of words for more prototypical concepts like *joy* may tend to resist change, more so than words for less prototypical ones like *optimism*; see Figure 1 for an illustration.

Although prototypicality has been discussed in other semantic categories, we do not expect prototypicality to predict semantic stability in every category. The basis of prototypicality and thus its ability to predict semantic change may differ in the classic example of birds [12]. The prototypicality of a bird name is primarily based on differences in biological taxonomies [13] and features grounded in sensory or visual perception [14]. As such, the features that define more (e.g., *sparrow*) or less prototypical birds (e.g., *penguin*) are equally well-defined, so the meanings of prototypical bird names do not help define the meanings of less prototypical bird names (see *Supplementary Information*), in contrast to the category of emotion words. We expect that while prototypicality should correlate with semantic stability of emotion words, it should not correlate with semantic stability of bird names.

Our hypothesis augments general principles of lexical evolution and semantic change. It has been shown that usage frequency is a general determiner of stability in English verb regularization [15], lexical replacement [16, 17, 18], loan word borrowing [19], and semantic change [20, 21]. If explained through the lens of communication, frequency should predict semantic stability: when speakers change the meaning of a highly frequent lexical item, they would face a higher number of misunderstandings than if they change a low-frequency item [22, 23]. As a result, we expect frequent emotion words to change less in meaning than other emotion words. We examine the prototypicality of emotion concepts as an additional predictor of semantic stability beyond usage frequency.

Our hypothesis differs from diachronic prototype semantics [24], which states that more prototypical senses of a word tend to stay prototypical over time and exhibit more stability than peripheral senses. Although this theory is consistent

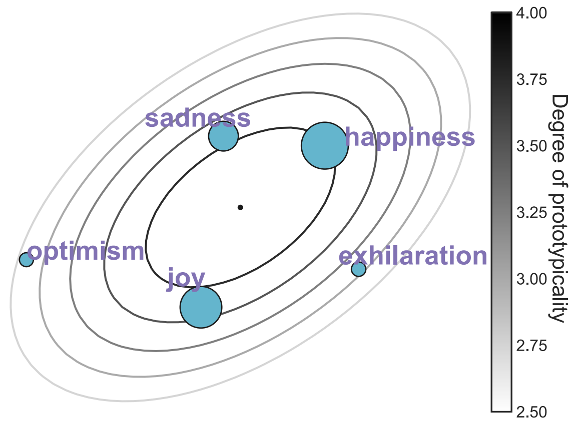

Figure 1: An illustration of the relation between prototypicality and semantic stability of emotion words. Each blue dot represents an emotion word, and the size of the dot is proportional to its predicted semantic stability; the smaller the dot, the higher its rate of semantic change over time. The contours indicate degrees of prototypicality. Visually, an emotion word close to the center has high prototypicality, and vice versa.

with our hypothesis regarding the pattern that prototypicality offers stability, we focus on explaining rates of semantic change among concepts in a lexical field, as opposed to characterizing principles of change among senses of an individual word [24, 25]. Previous studies have examined the theory of diachronic prototype semantics over the whole lexicon and found the prototypicality of words in statistical clusters (formed in meaning space) to negatively correlate with rates of semantic change [26, 21]. However, these studies do not explain how semantic change relates to prototypicality in the scope of a specific category such as emotions or birds.

We present a methodology for modeling emotion semantics and its evolution by building on work from machine learning and natural language processing in word embedding [27, 28, 29] and its historical extensions [30, 20, 31, 32]. We model emotion semantics using a vector-space representation trained on historical text corpora of natural language use, and we use this representation to model human judgments of prototypicality and semantic change of emotion words. Vector-space models of word meaning have been used within affective science for reconstructing human emotion ratings on dimensions such as valence and arousal [33], sentiment analysis [34], and analyzing emotion categories in documents [35], but not for investigating the open question of the evolution of emotion semantics.

2. Methodologies for quantifying rates of semantic change

Quantifying the rate of semantic change for a word requires records of its meaning from two distinct time periods and a quantitative metric that compares these records. One type of methods that constructs word meanings and enables comparisons over time is based on word embeddings [27, 28]. The embedding of a word is a real-valued vector that represents its meaning through a high-dimensional space; vectors for words with similar meanings tend to be close in this space, such as *compassion* and *sympathy*. Word embeddings are constructed from co-occurrence statistics in large text corpora. We thus obtain meaning representations from two distinct time periods by constructing word embeddings based on historical text corpora from the corresponding periods [20].

Existing methods for computing rates of semantic change often rely on the cosine distance between two embeddings [20, 21]. According to this metric, a large cosine distance between historical and contemporary embeddings implies a high rate of semantic change, and vice versa. However, this metric tends to bias the correlation between rate of semantic change and frequency [21]. For this reason, we use an alternate neighbourhood-based metric to compare word embeddings across time [36]. This metric quantifies the rate of semantic change for a word w between periods t_1 and t_2 via the Jaccard distance between sets of k-nearest neighbours in meaning space:

$$rate(w, t_1, t_2) = 1 - \frac{|kNN(w, t_1) \cap kNN(w, t_2)|}{|kNN(w, t_1) \cup kNN(w, t_2)|} \tag{1}$$

where $kNN(w, t)$ contains the k words whose embeddings are the closest to the embedding of w in terms of cosine similarity. Intuitively, we say a word underwent semantic change if the composition of its semantic neighbourhood has changed. Following [36], the part of speech (POS) of the members of $kNN(w, t)$ is always the same as the POS of w, and we also set k to 100. In *Supplementary Information*, we show that this measure is robust to variations in k. Compared to the cosine metric, this metric enables more transparent interpretation on rates of change because we can inspect and evaluate the sets of semantic neighbours (see *Supplementary Information* for examples of emotion semantic change).

To implement this metric at scale, we used pretrained historical word embeddings and POS tags from HistWords [20]. Specifically, we used 300-dimensional Word2Vec (SGNS) embeddings obtained from the Skip-Gram model [28] and trained on the corpora Google Books Ngrams English and French. We used historically most frequent POS tags from the same sources. This provided us with historical word embeddings and most frequent POS tags for 100,000 English words and 100,000 French words, for every decade between 1800 and 2000.

3. Analyses of emotion concepts

In the first set of analyses, we provide evidence for our hypothesis that the well-defined features and functions of prototypical emotion words promote semantic stability. Specifically, we test against the null hypothesis that prototypicality does not predict semantic stability in English and French emotion words over the past century.[2] We describe resources that provide us with lists of English and French emotion words, emotion prototypicality ratings, and historical frequency estimates. We then describe our methods for estimating prototypicality ratings historically and for hypothesis testing, which is followed by a presentation of our results.

[2] We focused on these two languages because 1) we want to test if our analysis generalizes beyond a single language, and 2) there is a limited cross-linguistic variety of empirical studies on emotion prototypicality and of the historical data provided by HistWords.

3.1. Materials

We obtained a list of English emotion words and their corresponding prototypicality ratings from [6]. The authors produced the list by obtaining 213 emotion nouns from a collection of emotion concepts. They produced emotion prototypicality ratings by asking 112 American university students to rate each of these nouns on a 4-point scale, where 4 means the noun is definitely an emotion, and 1 means the noun is definitely not an emotion. Following this work, our analyses focused on nouns that have prototypicality ratings at least 2.75 with the addition of *surprise* and exclusion of *abhorrence*, *ire*, *malevolence*, and *titillation*; we additionally included the word *awe*. We also obtained the valence of these emotion words from the study, which was originally derived from applying multidimensional scaling to similarity judgments [6].

We also obtained a list of French emotion words with their corresponding prototypicality ratings [37]. The authors produced the list by translating 237 Italian emotion words from an earlier study into French. They produced emotion prototypicality ratings by asking 319 French university students to rate each of these words on a 10-point scale, where 10 means the word is certainly an emotion, and 1 means it is not an emotion. To be consistent with the English list, we kept emotion words whose most frequent POS tag is noun in the final decade of our historical POS data. We also obtained the valence of these emotion words from the study, which was originally obtained by asking 300 French university students to rate the words on a scale of -5 (very unpleasant) to 5 (very pleasant) [37].

We obtained historical frequency data from HistWords [20], which is based on the corpora Google Books Ngrams English and French. This yielded historical frequency data for 682,459 English words and 213,686 French words, for every decade between 1800 and 2000. We intersected the word lists with historical word embeddings, POS tags, and frequency from HistWords. We noticed that more emotion words were unavailable when we increased the span between flanking decades than otherwise: if $t_1 = 1890$ and $t_2 = 1990$, only 9 words from the English list are unavailable in HistWords and the HTE, but if we used $t_1 = 1800$, the number increased to 28; similarly in French, the shorter time span resulted in 32 unavailable words, but the longer one resulted in 58 unavailable words. Consequently, we decided to use the decades of 1890 and 1990 as the flanking decades for our analysis (i.e. $t_1 = 1890$, $t_2 = 1990$), and we used historical frequency data from the 1890s. After the intersection, we had a total of 123 English emotion words and 112 French emotion words.

3.2. Methods

Since we cannot go back in time to measure the prototypicality of emotion concepts in the past, we needed a method for estimating historical prototypicality. Let x represent the word embedding of a concept in category c. Following previous work in prototype theory [38, 39], we estimated the prototypicality of x as the unnormalized conditional probability $p(c|x)$, which can be computed using an isotropic Gaussian via Bayes rule:

$$p(c|x) \propto p(x|c) \sim N(\mu, I) \tag{2}$$

where $\mu = \frac{1}{|E_c|} \sum_{v \in E_c} v$ and E_c is the set of embeddings for members of c; I is an identity matrix. Intuitively, we estimated the prototypicality of x by computing its distance from the category centroid μ; the closer they are, the higher its estimated prototypicality. To estimate the prototypicality of an emotion concept in history, we used its historical embedding x and the embeddings of other emotion concepts to compute $p(x|c = emotion)$. We evaluated this method by computing the correlation between our empirical prototypicality ratings obtained from [6, 37] and our estimated prototypicality based on embeddings from the 1980s and 1990s, the decades closest to the publication of those studies.

To test against the null hypothesis, we computed the rate of change for every emotion concept x, $rate(x, 1890, 1990)$ using Equation (1) and historical embeddings and POS tags from HistWords. Separately for English and French, we then computed the Pearson correlations between the emotion concepts' rates of change and prototypicality estimated for the 1890s. To evaluate whether the prototypicality of emotion concepts predicts rates of change beyond frequency, we performed multiple linear regressions for English and French using the following regression formula:

$$rate(x, 1890, 1990) \sim p(x|c = emotion) + freq(x) + val(x) \quad (3)$$

where for every concept x, we denote its usage frequency as $freq(x)$ and its valence as $val(x)$, which we added to control for unequal numbers of negative and positive emotion concepts in our datasets. We fitted the model using ordinary least squares implemented by `statsmodel` [40]; we also used this package to compute relevant test statistics. Following previous work [20], we performed a log transformation on frequency.

3.3. Results

Figure 2 shows the Pearson correlation between estimated prototypicality from English word embeddings and ratings from English speakers [6]: $\rho = 0.428$, $p < 0.001$, $n = 123$. We obtained similar results with French word embeddings for a set of French emotion concepts [37]: $\rho = 0.438$, $p < 0.001$, $n = 112$. These initial results show our estimated degrees of prototypicality for emotion concepts capture human judgments to some extent. For this reason, we used the same method to estimate historical prototypicality ratings and evaluated them as predictors of semantic stability.

Figure 3 shows a significant negative correlation between emotion prototypicality and degree of semantic change: $\rho = -0.580$, $p < 0.001$, $n = 123$. On average, emotion concepts rated prototypical such as *anger*, *joy*, *fear* underwent less change in meaning compared to words denoting less prototypical concepts such as *zest, exhilaration* and *hysteria* (see annotated word samples in Figure 3). Similar results hold for French: $\rho = -0.576$, $p < 0.001$, $n = 112$. *Supplementary Information* provides additional examples of English and French emotion concepts from the most changing to the most semantically stable, along with their semantic neighbours retrieved from our methods.

Figure 4 shows our results for multiple regression. The adjusted R^2 of the model for English is 0.680, with $p < 0.001$, $n = 123$; mean regression coeffi-

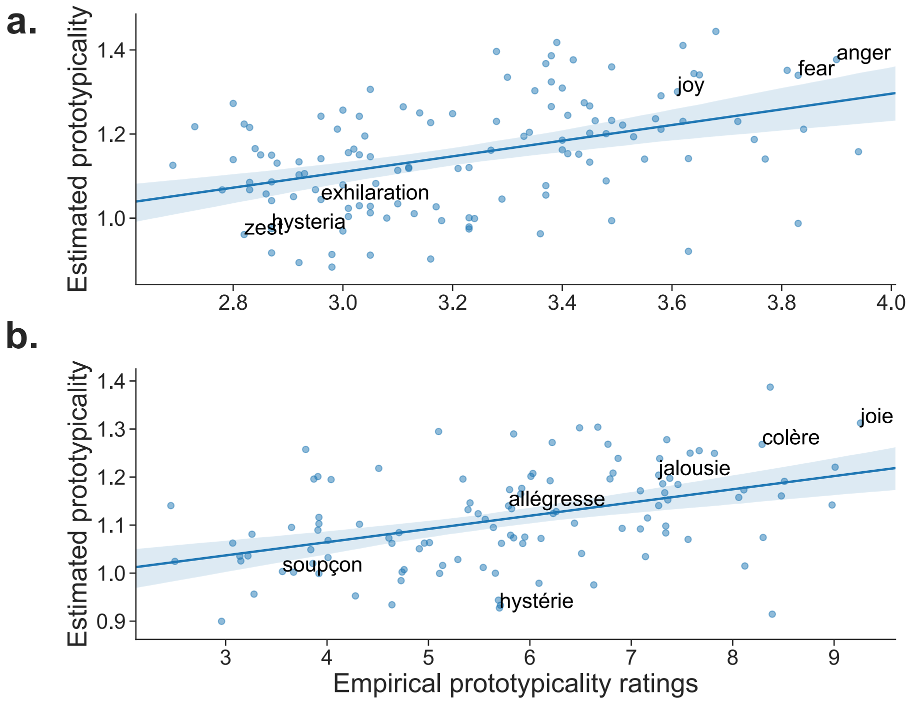

Figure 2: Word embedding reconstruction of emotion prototypicality in a) English and b) French. Scatter plots compare estimated prototypicality computed from Equation 2 against empirical ratings. Each dot corresponds to an emotion concept (a sample of concepts annotated), and each band shows a 95% confidence interval for the line of best fit.

cients for prototypicality ($\beta = -0.417$, $p < 0.001$) and frequency ($\beta = -0.0451$, $p < 0.001$) are significant, but for valence ($\beta = 0.0053$, $p = 0.208$) it is insignificant. For French, the adjusted R^2 of the model is 0.538, with $p < 0.001$, $n = 112$; mean regression coefficients for prototypicality ($\beta = -0.6363$, $p < 0.001$) and frequency ($\beta = -0.0331$, $p < 0.001$) are significant, but for valence ($\beta = 0.0019$, $p = 0.454$) it is insignificant. These results show that frequency predicts semantic stability, which confirms the previous findings [20, 21]. Beyond frequency, we find that prototypicality plays an important role in predicting semantic stability of emotion words, manifested in its significant and negative effect. This provides evidence for our hypothesis that prototypical emotion words tend to

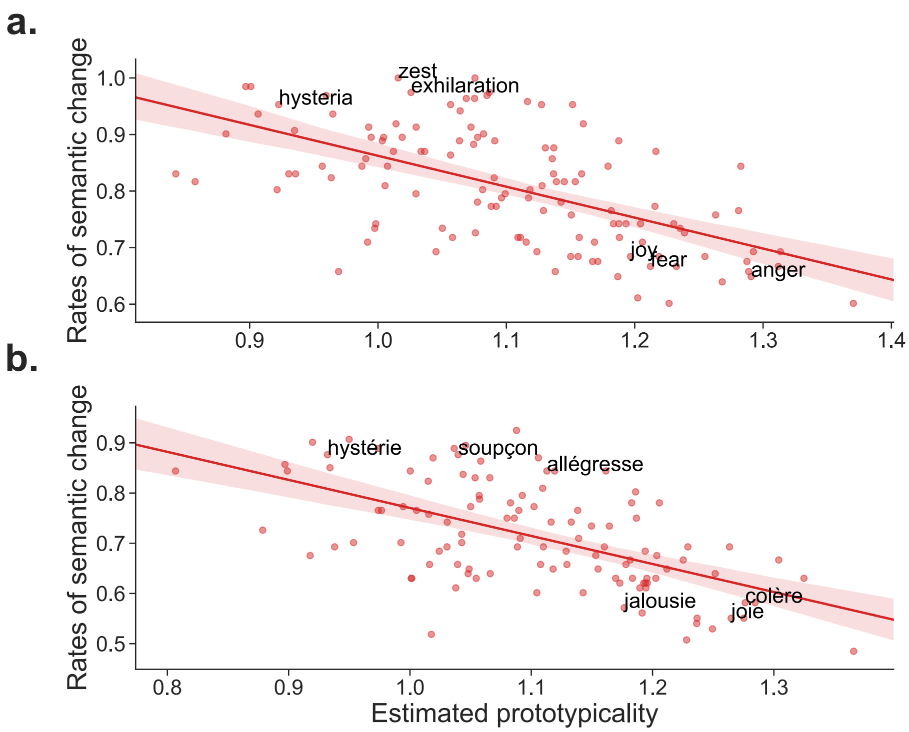

Figure 3: Scatter plots showing the negative correlations between emotion prototypicality and rates of emotion semantic change between the 1890s and 1990s, in a) English and b) French. Each dot corresponds to an emotion word (with a sample set of words annotated), and each band shows a 95% confidence interval for regressions between emotion prototypicality and rates of semantic change.

be semantically stable over time.

Supplementary Information includes three more analyses that further corroborate our findings. The first analysis repeats the multiple regression but restricts the neighbourhoods to emotion concepts only when computing $rate(w, 1890, 1990)$; the results rule out the possibility that our findings are an artifact of the non-emotion senses of polysemous emotion concepts (e.g., *zest*). The second analysis extends the multiple regression for English by including additional predictors based on hypernymy-hyponymy, age of acquisition, and degrees of polysemy, which could potentially subsume the effects of prototypicality; our results

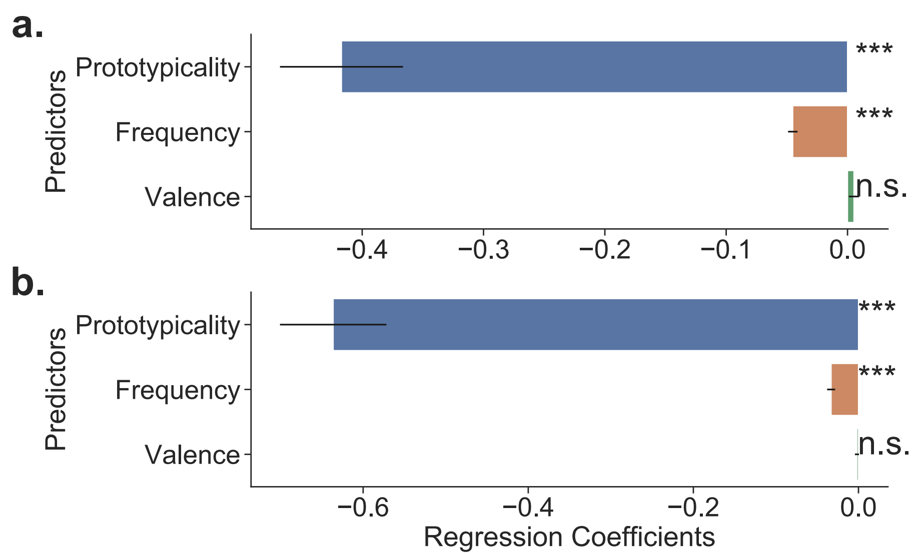

Figure 4: Predictor coefficients from multiple regressions on rates of emotion semantic change. Error bars show standard error, and "n.s.", "*", "**", "***" denote no significance at $p < 0.05$, and $p < 0.05$, $p < 0.01$, $p < 0.001$ respectively. a) shows results for English, and b) shows results for French.

show that this is not the case. The third analysis repeats the multiple regression for English emotion concepts, except the rates of change are computed as $rate(w, 1980, 1990)$ and empirical prototypicality from [6] were used; these results provide evidence that the effect of prototypicality is not caused by potential artifacts in our estimation of prototypicality based on Equation 2.

Figure 5 illustrates our main finding with two example words: *disgust* and *awe*. These words had similar usage frequencies over time, but *disgust* is rated as a more prototypical emotion word than *awe* [6]. Over time, *awe* has shifted meaning more substantially than *disgust*. In particular, both words were in the neighbourhood of negative emotion words (e.g., *sadness*, *anger*, and *fear*) in the 1890s. However, while *disgust* still remained close to these words in the 1990s, *awe* moved closer to positive emotion words (e.g., *love* and *happiness*).

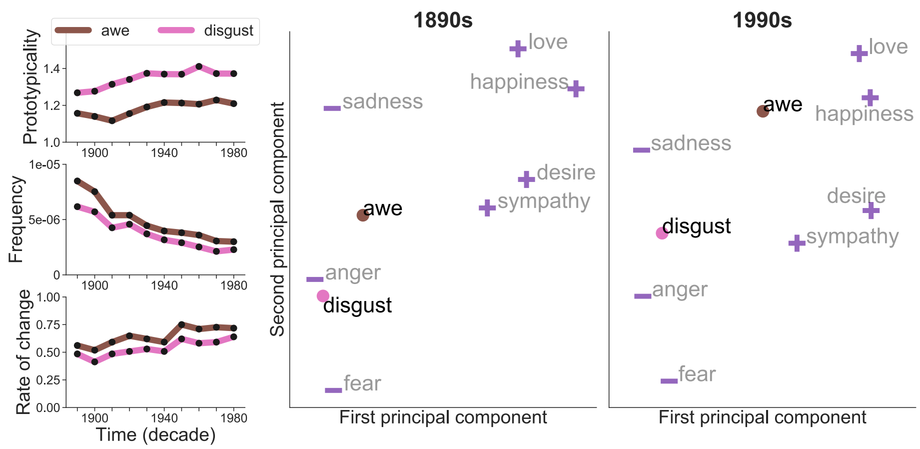

Figure 5: An illustrative comparison of prototypicality, frequency, and semantic stability in emotion words *awe* and *disgust*. Left panels show the embedding-based prototypicality, frequency and degree of semantic change of *awe* and *disgust* over time. Right panels visualize the rates of change in these words by placing them in the two principal components [41] of meaning space, along side prototypical emotion concepts which are annotated based on their valence ("+" for positive, "-" for negative).

4. Analyses of bird names

In this set of analyses, we demonstrate that the sources of prototypicality do not always provide semantic stability as we have shown for emotion concepts. Here we repeat our previous analyses on a case study of birds, a frequently investigated category in prototype theory [12, 42]. As we will see, our embedding-based estimation of prototypicality does not work well with bird names, and we will focus our analysis on using empirical ratings from the 1970s.

4.1. Materials

We obtained a list of English bird names with prototypicality ratings from [12]. The author produced the list by consulting previous work so that the selected names were relatively frequent. They produced bird prototypicality ratings by asking 209 American university students to rate each of these names on a 7-point scale, where 1 means the name refers to a very good example of a bird, and 7 means the name refers to a very poor example. Note that the scale operates in the opposite direction of our prototypicality ratings for emotion concepts. For clarity, we multiplied these ratings by -1 so the direction is the same as our emotion data. Focusing on the 1970s and 1990s, we used historical data from HistWords [20], which was intersected with the word list and provided us with 41 bird names.

4.2. Methods

Similar to the previous section, we attempted at estimating bird prototypicality using Equation 2. We then computed the rates of change for every bird name w, $rate(w, 1970, 1990)$ using Equation 1. We computed the Pearson correlation between rates of change and prototypicality ratings obtained from the 1970s, and we performed a multiple regression using the following formula:

$$rate(w, 1970, 1990) \sim proto(w) + freq(w) \tag{4}$$

where we denote the empirical prototypicality rating of every bird name w as $proto(w)$.

4.3. Results

Figure 6a shows the Pearson correlation between estimated prototypicality and empirical ratings from [12]: $\rho = 0.153$, $p = 0.340$, $n = 41$. While the same method reconstructs prototypicality for emotion concepts to some extent, our text-based method does not explain a significant amount of variance in the prototypicality of birds which depends more on sensory features [14]. It has been shown that prototypical birds in our dataset tend to be passerines, small perching birds that sing (e.g., *robin*), and less prototypical ones tend to be non-passerines (e.g., *penguin*) [13], which our text-based methodology did not capture. For this reason, we chose to focus on empirical prototypicality ratings for birds in our analyses.

Figure 6b shows a significant positive correlation between bird prototypicality and degree of semantic change: $\rho = 0.428$, $p = 0.005$, $n = 41$. This finding suggests that the relation between semantic change and prototypicality in bird names is opposite to our previous findings for emotion words. Figure 7a shows the results for multiple regression. The adjusted R^2 is 0.508, with $p < 0.001$, $n = 41$; mean regression coefficients for empirical prototypicality ($\beta = 0.0283$, $p = 0.011$) and frequency ($\beta = -0.0454$, $p < 0.001$) are significant. We observe frequency still predicts semantic stability, suggesting it is indeed a general predictor of semantic change. Interestingly, prototypicality of birds not only failed to predict stability as in the case of emotion concepts, but also pointed to the opposite trend: in the category of birds, names of prototypical birds tend to undergo more change than other names.

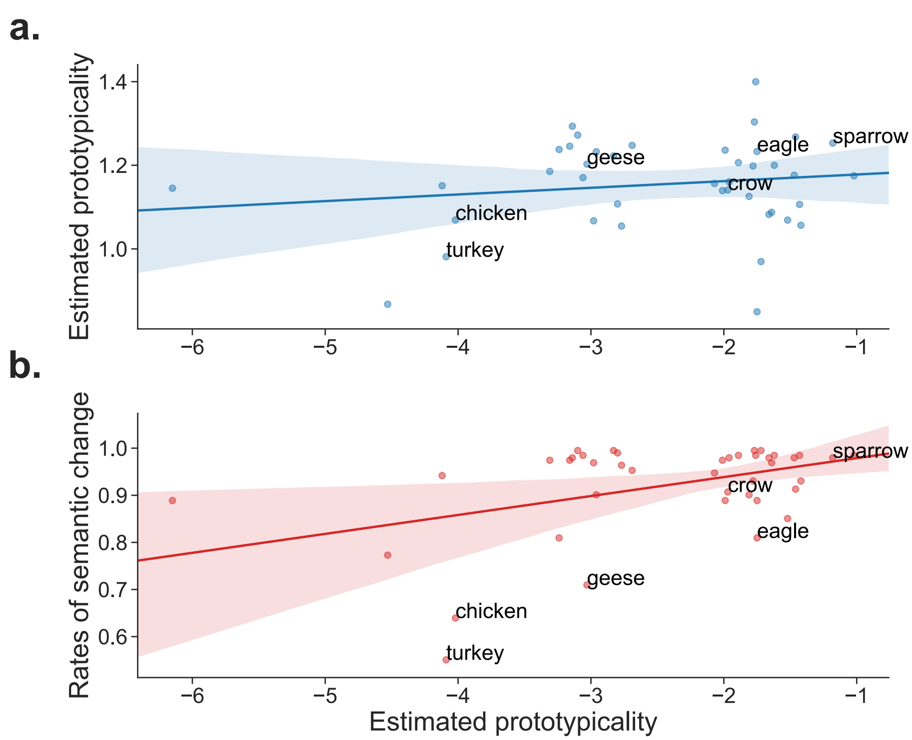

Figure 6: Analyses of bird names: a) word embedding reconstruction of bird prototypicality and b) correlations between bird prototypicality and rates of semantic change between the 1970s and 1990s. Each dot corresponds to a bird name, and each band shows a 95% confidence interval for the line of best fit.

To better understand the implications of this variability to our finding about

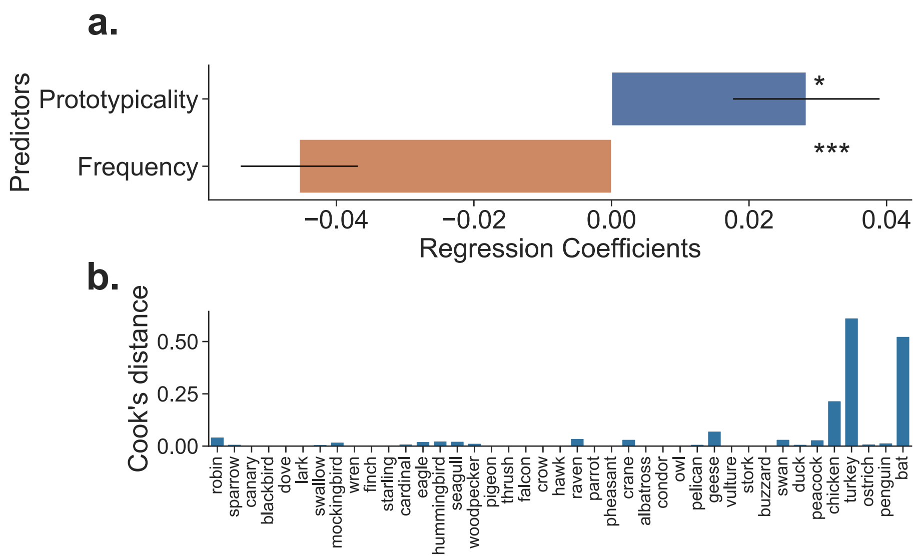

Figure 7: Multiple regression analysis of bird names: a) predictor coefficients from multiple regressions on rates of semantic change, following the same layout as Figure 4; b) Cook's distance for every bird name, showing the influence of individual data points on the regression result.

bird names, we performed a more in-depth analysis of the data. Unlike the case of emotion, we observe bird names exhibit high variability in Figure 6b, which is reflected in the wide confidence region.[3] This suggests the opposite trend in bird names is influenced by only a handful of less prototypical birds. We estimated the influence of each bird name using Cook's distance, which takes into account the data point's residual and leverage. Figure 7b shows the influence of each bird in the regression analysis: the most influential points are *turkey*, *bat* and *chicken*. We can observe *bat* is likely to be influential as it has a much higher rating (not prototypical) than other bird names; this might be because subjects in the original study, being university students, were familiar with the scientific classification of bats. More importantly, despite

[3]Note that the number of available bird terms for our analysis is substantially lower than that of emotion terms.

not being prototypical birds, *turkey* and *chicken* could have important cultural roles (festive or culinary) in North America so that they provided anchors for their meaning, thereby contributing to the significant correlation between bird prototypicality and semantic change.

Figure 8 compares the degrees of semantic change that took place in emotion concepts and bird names between the 1970s and 1990s. Many prototypical emotion concepts tend to lie at the lower tail of the density distribution and show high stability, mirroring the results we have seen previously, but the same pattern does not hold for birds. We observe that overall bird names tend to undergo greater change than emotion concepts do. It is possible that prototypical birds possess the most representative features of the bird category, which could provide points of attachment for meaning change via processes such as chaining, in which a word for one object is extended to be used for another, or metaphor [43, 25]. This general pattern of more rapid change among bird names together with the additional semantic stability of a handful of influential bird exemplars may be responsible for the positive correlation between degrees of bird prototypicality and rates of semantic change.

5. Conclusion

Language offers a lens into the history of emotion semantics. Our computational linguistic analyses of semantic change suggest that a new view of emotion concepts in language may be warranted. Rather than perceiving of emotion concepts as static, their meaning is evolving over time. The exact cultural or societal factors responsible for semantic change in emotion words are difficult to pinpoint, and they may be different for each emotion term. For example, semantic change in *awe* may reflect a movement away from its use in religious contexts, in which it reflects more of a fearful respect, towards greater use in beautiful artistic and natural contexts that followed the emergence of romanticism and transcendentalism in the early to middle nineteenth century. We assessed semantic change over a relatively short timescale, suggesting that in

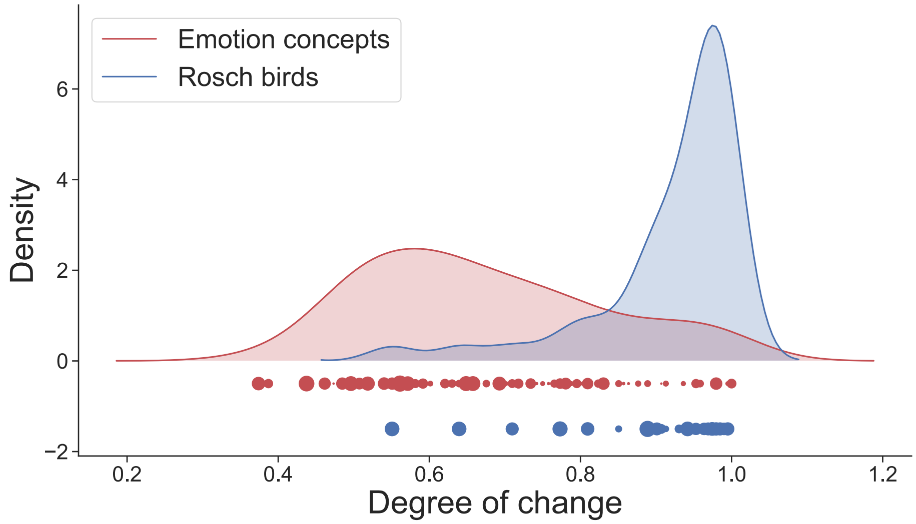

Figure 8: Distributions of semantic change in emotion and bird categories. Each dot corresponds to a word, and the size of the dot is proportional to its degree of prototypicality. The density plots were obtained using kernel density estimation; although degrees of change given by Equation 1 are technically bounded between 0 and 1, we did not bound the support of this figure for illustrative reasons.

the centuries to come it is possible that words like *awe* may continue to evolve and mean something very different than they do today.

Further, we found in two languages that more prototypical emotion words [6, 42] showed greater semantic stability than other emotion words over time. The relation between prototypicality and semantic change depends on its exact sources, as we observed opposite trends for emotions and birds. The importance of prototypicality as a predictor in semantic change for other semantic categories remains an open question and future work should investigate what features affect the importance of prototypicality. Our study extends research on emotions to its historical development and offers a computational cognitive characterization of evolving emotion semantics from natural language use.

Acknowledgements

We thank Brett Ford and two anonymous reviewers for providing constructive comments on our manuscript. This work was supported by a University of Toronto Graduate Entrance Scholarship to AX, a Canada Foundation for Innovation grant #502426 to JES, a NSERC Discovery Grant RGPIN-2018-05872 and a SSHRC Insight Grant #435190272 to YX.

Appendix A. Supplementary material

Code and data used for our analyses are available on GitHub at `https://github.com/johnaot/Emotion_Semantic_Change`.

Appendix B. Word age and prototypicality

Here we further demonstrate the differences between prototypicality in emotions and prototypicality in birds. Previous work has suggested that prototypical emotion concepts are well-defined and may have particularly strong social or cultural scripts [8, 5, 9]. Thus it is conceivable that words for prototypical emotion concepts exist in a lexicon prior to those less prototypical emotion words. In contrast, prototypicality of birds is based on biological taxonomy [13] and grounded in sensory and visual perception [14], and we do not expect prototypicality to be reflected in the age of a word; it is likely that a relatively newly documented passerine (e.g., bluebird) entered the lexicon after well-established non-passerines (e.g., chicken). To test these ideas, we analyzed the correlation between word age and prototypicality in the categories of emotion and birds.

Following existing work [44], we obtained the age of a word from the Historical Thesaurus of English (HTE) [45]. For each word entry, the HTE provides a list of senses of the word, and for each sense, the HTE provides the word class associated with that sense of the word and the date of first appearance of the sense in historical records. We operationalized the date of (the first) emergence of a word to be the earliest date among the dates of first appearance across all of its senses. Since our analyses focused on nouns, we considered only noun senses. We did not analyze the age of French words due to the unavailability of comparable French dictionaries.

We analyzed the same lists of English emotion words and bird names described in *Section 3.1* and *Section 4.1*, which we intersected with the HTE data. For English emotion words, the Pearson correlation between emerging date and prototypicality is -0.366, $p < 0.001$, $n = 135$. This indicates that prototypical emotion words emerged earlier in the history of English. The same pattern does not hold for the bird category. For Rosch bird names, the Pearson correlation between emergence date and prototypicality is -0.0657, $p = 0.643$, $n = 52$. See Figure B.9 for illustration.

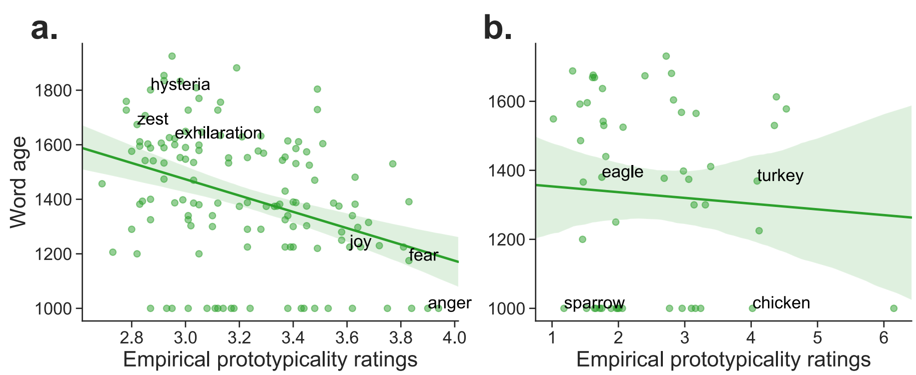

Figure B.9: Scatter plots showing the relations between prototypicality and word age for a) emotion words and b) bird names. Each dot corresponds to a word (with a sample set of words annotated), and each band shows a 95% confidence interval for regressions between prototypicality and age.

Appendix C. Evaluation of the nearest-neighbour measure

We show that the nearest-neighbour measure of semantic change described in *Section 2* is 1) robust to variation in neighbourhood size (denoted by k), and 2) interpretive for a word's semantic change based on nearest neighbours retrieved at different time points.

Appendix C.1. Robustness in neighbourhood size k

We evaluate whether the nearest-neighbour measure is robust to variation in k. Following *Section 3.1*, we again quantify semantic change by setting $t_1 = 1890$ and $t_2 = 1990$ and focus our analysis on the lists of English and French emotion words we analyzed in the main text. We show that for $k = 20, 40, 60, 80$, and 100, resulting degrees of semantic change $rate(w, t_1, t_2)$ are highly correlated. The correlation results are summarized in Figure C.10. We observed that in both English and French, the degrees of change are strongly correlated between any two of the predetermined settings of k, with a small decrease in correlations for the lowest value of $k = 20$.

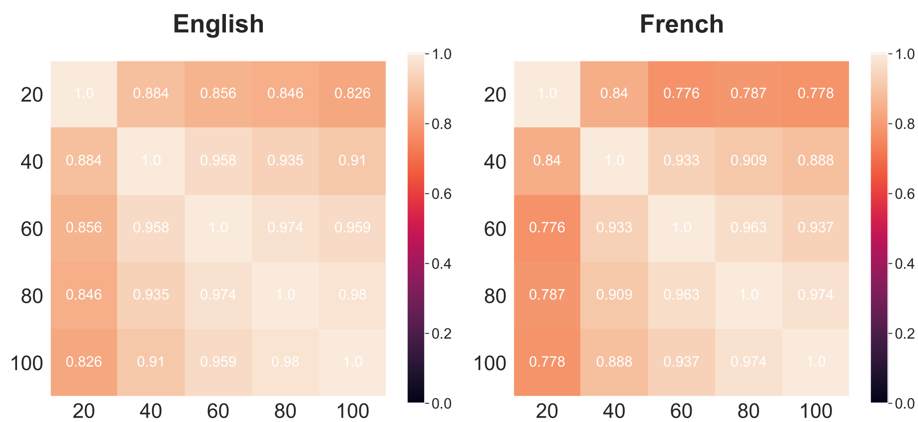

Figure C.10: Robustness of kNN demonstrated using emotion words. The first column shows results for English and the second column shows results for French. Each cell shows the Pearson correlation between changes measured by x-nearest neighbours and by y-nearest neighbours. All p-values are significant and less than 0.001.

Appendix C.2. Examples of emotion semantic change

Qualitative changes in the nearest neighbours of a word offer interpretability for semantic change. We provide examples of emotion words that underwent the most and the least changes and their nearest semantic neighbours in Tables C.1, C.2, C.3, and C.4. For example, in Table C.1, we observed *zest*, which used to primarily convey joy but later became primarily associated with food, is among the most changing emotion words; on the other hand, in Table C.2, we observed words like *surprise* has barely changed in meaning. Similarly, in Table C.3, we observed *stupéfaction* in French become less associated with despair and anxiety over time; in Table C.4, the least changing words on the other hand tend to preserve similar emotion words as neighbours.

Most Changing	Nearest Neighbours in 1890s	Nearest Neighbours in 1990s
zest	relish, enjoyment, sprightliness	juice, teaspoons, vinegar
infatuation	priestcraft, devastations, misanthrope	inhomogeneity, palates, pleurisy
sentimentality	cant, sentimentalism, rusticity	polyphony, sterne, mandel
optimism	pessimism, aptness, sentimentalism	pessimism, insecurity, enthusiasm
exhilaration	mountebank, festivity, tulip	joy, sadness, excitement
aggravation	misery, symptoms, consequences	proxies, sleeplessness, stressor
exasperation	vehemence, peevishness, pitch	astonishment, amazement, disgust
glee	merriment, wassail, delight	claps, shouts, megaphone
cheerfulness	hopefulness, sprightliness, vivacity	serenity, blasphemies, cannonade
gaiety	sprightliness, vivacity, gayety	pontellier, faints, plaudits
fondness	liking, peevishness, passion	affection, pianists, groanings
hysteria	neurasthenia, simulation, melancholia	neurosis, hypochondriasis, psychosis
dejection	despondency, sullenness, irresolution	pantomime, theseus, disquiet
elation	sullenness, despondency, peevishness	despair, revulsion, dread
ferocity	fierceness, cruelty, prowess	vigor, fury, proverb
revulsion	feeling, disquietude, outburst	disgust, hisses, yearnings
isolation	loneliness, disorganization, seclusion	monger, characterization, coli
alienation	eviction, property, repugnancy	helplessness, blauner, resentment
hopelessness	uselessness, futility, helplessness	helplessness, despair, frustration
rapture	ecstasy, delight, joy	joy, indignation, outcasts

Table C.1: Top 20 most changing English emotion words along with their 3 nearest neighbours in the flanking decades.

Appendix D. Additional analyses of emotion semantic change

We describe three additional analyses that corroborate our findings on emotion semantic change in the main text. The first analysis rules out the possibility that our findings are an artifact of the non-emotion senses of polysemous emotion concepts (e.g., *zest*). The second analysis shows that additional predictors based on hypernymy-hyponymy and degrees of polysemy do not subsume the

Least Changing	Nearest Neighbours in 1890s	Nearest Neighbours in 1990s
grief	sorrow, anguish, joy	sorrow, sadness, anguish
pity	compassion, love, sympathy	compassion, shame, sadness
misery	wretchedness, miseries, degradation	sorrow, bitterness, anguish
disgust	horror, aversion, indignation	sadness, annoyance, amazement
anger	indignation, resentment, rage	resentment, rage, frustration
surprise	astonishment, amazement, dismay	astonishment, amazement, dismay
sorrow	grief, anguish, sadness	grief, sadness, misery
affection	affections, tenderness, esteem	admiration, sympathy, love
happiness	felicity, prosperity, welfare	prosperity, joy, enjoyment
despair	desperation, dismay, rage	anguish, frustration, sadness
fear	dread, anger, shame	dread, resentment, anger
horror	terror, astonishment, amazement	terror, astonishment, amazement
regret	disappointment, grief, sorrow	disappointment, sadness, bitterness
envy	jealousy, uncharitableness, hatred	jealousy, hatred, resentment
disappointment	mortification, grief, regret	frustration, sadness, regret
rage	fury, anger, indignation	anger, fury, indignation
shame	disgrace, infamy, blush	guilt, pity, humiliation
astonishment	amazement, surprise, dismay	amazement, dismay, surprise
joy	gladness, delight, grief	delight, sorrow, excitement
sympathy	sympathies, compassion, affection	affection, admiration, compassion

Table C.2: Top 20 least changing English emotion words along with their 3 nearest neighbours in the flanking decades.

effects of prototypicality on emotion semantic change. The third analysis provides evidence that our results on emotion concepts were not caused by artifacts in our estimation of prototypicality.

Appendix D.1. Category-bounded analysis

We investigate the robustness of our semantic change analyses by considering a variant of the nearest-neighbour measure discussed in *Section 2*. Originally in *Section 2*, the degree of semantic change of a word w is defined as the Jaccard distance between its nearest neighbours at time t_1, $kNN(t_1)$, and its nearest neighbours at a later time, $kNN(t_2)$, where kNN is restricted to nouns in the entire lexicon and determined by cosine similarity over word vectors. Since the meaning of every emotion word is represented by one word vector only, the set kNN might also capture meaning change with respect to the word's polysemes

Most Changing	Nearest Neighbours in 1890s	Nearest Neighbours in 1990s
stupéfaction	indiscrétion, désespoir, anxiété	désapprobation, émotion, allégresse
suspicion	inculpation, défiance, prévention	méfiance, défiance, incertitude
culpabilité	réussite, identité, présomption	infériorité, persécution, châtiment
déplaisir	étonnement, inquiétude, appréhensions	plaisir, mâle, océans
torpeur	engourdissement, apathie, léthargie	apathie, consternation, stupeur
extase	contemplation, stupeur, somnambulisme	contemplation, joie, angoisse
soupçon	soupçons, équivoque, délit	préjugé, partialité, préjugés
hystérie	épilepsie, diabète, étiologie	névrose, épilepsie, névroses
séduction	adultère, entraînements, cruauté	immédiateté, impiété, éloquence
désolation	épouvante, dévastation, misère	pauvreté, saleté, nausées
froideur	bonhomie, trousseau, bassesse	arrogance, ingratitude, insolence
excitation	irritation, nerfs, nerf	tension, angoisse, agitation
intimidation	violence, ruse, corruption	menaces, coercition, chantage
timidité	fierté, délicatesse, naïveté	docilité, avidité, découragement
tension	volts, pression, potentiel	appareillage, excitation, tensions
intérêt	intérêts, utilité, équité	intérêts, utilité, rentabilité
espérance	espoir, espérances, désir	espoir, espérances, désir
dépit	mépris, défaillances, hésitations	grâce, précocité, conséquence
allégresse	fierté, épouvante, gaieté	joie, émotion, gaieté
effusion	sang, tendresse, larmes	sang, amertume, lucre

Table C.3: Top 20 most changing French emotion words along with their 3 nearest neighbours in the flanking decades.

and homonyms, i.e., meaning change outside the category of emotion. To assess how such meaning change might affect our results, we restricted the set of nearest neighbours so that only the list of emotion words are included, i.e., a category-bounded analysis of emotion semantic change. Since the set of emotion words is much smaller than the full lexicon, we set the size of the neighbourhood to be $k = 25$.

We first provide evidence that this variant of the nearest-neighbour measure is also capable of capturing semantic change by showing 1) this measure is positively correlated with the original nearest-neighbour measure, and 2) this measure captures the negative relationship between frequency and semantic change [20]. We obtain degrees of change under this variant measure by following the same procedure described in *Section 2*. In the case of English emotion words, the Pearson correlation between degrees of semantic change measured by

Least Changing	Nearest Neighbours in 1890s	Nearest Neighbours in 1990s
tristesse	angoisse, amertume, effroi	amertume, douleur, angoisse
tendresse	bienveillance, sympathie, sollicitude	douceur, compassion, amour
patience	courage, persévérance, prudence	courage, persévérance, audace
orgueil	vanité, amour, ambition	arrogance, insolence, vanité
horreur	honte, effroi, angoisse	opprobre, effroi, tristesse
effroi	tristesse, consternation, terreur	tristesse, consternation, horreur
indignation	admiration, effroi, cri	enthousiasme, cri, admiration
joie	tristesse, douleur, bonheur	tristesse, enthousiasme, douleur
honte	horreur, chagrin, humiliation	opprobre, humiliation, peur
jalousie	haine, ambition, convoitise	arrogance, haine, rancune
colère	désespoir, mécontentement, anxiété	désespoir, fureur, émotion
douleur	chagrin, souffrance, douleurs	souffrance, tristesse, douleurs
stupeur	tristesse, angoisse, effroi	consternation, surprise, effroi
vengeance	haine, ressentiment, fureur	haine, jalousie, orgueil
bonheur	malheur, gloire, joie	malheur, joie, plaisir
chagrin	douleur, tristesse, honte	tristesse, douleur, amertume
terreur	effroi, horreur, haine	horreur, effroi, anarchie
enthousiasme	ardeur, joie, admiration	joie, admiration, indignation
impatience	anxiété, indignation, angoisse	joie, humeur, espoir
souffrance	douleur, tristesse, angoisse	douleur, angoisse, souffrances

Table C.4: Top 20 least changing French emotion words along with their 3 nearest neighbours in the flanking decades.

this variant and degrees obtained by the original measure is 0.751, $p < 0.001$, $n = 123$; the Pearson correlation between degrees of change measured by the variant and frequency is -0.489, $p < 0.001$, $n = 123$. In the case of French emotion words, the Pearson correlation between degrees of change measured by this variant and degrees obtained by the original measure is 0.604, $p < 0.001$, $n = 112$; the Pearson correlation between degrees of change measured by the variant and frequency is -0.203, $p = 0.0318$, $n = 112$. These results suggest this variant is capable of replicating patterns of change identified previously.

After validating this variant measure, we also repeated the analyses on emotion semantic change described in the main text. Figure D.11 shows a significant negative correlation between prototypicality and degree of semantic change: for English, $\rho = -0.535$, $p < 0.001$, $n = 123$; for French, $\rho = -0.558$, $p = 0.002$, $n = 112$. Figure D.12 shows multiple regression results similar to the results presented in the main text. The adjusted R^2 for English is 0.432, with $p < 0.001$, $n = 123$. Mean regression coefficients for prototypicality ($\beta = -0.479$, $p < 0.001$) and frequency ($\beta = -0.0356$, $p < 0.001$) remained negative and significant, whereas valence ($\beta = 0.0112$, $p = 0.091$) is insignificant. Again, results hold similarly for French with the adjusted $R^2 = 0.381$, $p < 0.001$, $n = 112$ (prototypicality $\beta = -0.600$, $p < 0.001$; frequency $\beta = -0.0208$, $p < 0.001$; valence $\beta = -0.0015$, $p = 0.599$). Compared to the main results, we observed that prototypicality remains a competitive predictor of semantic stability relative to frequency.

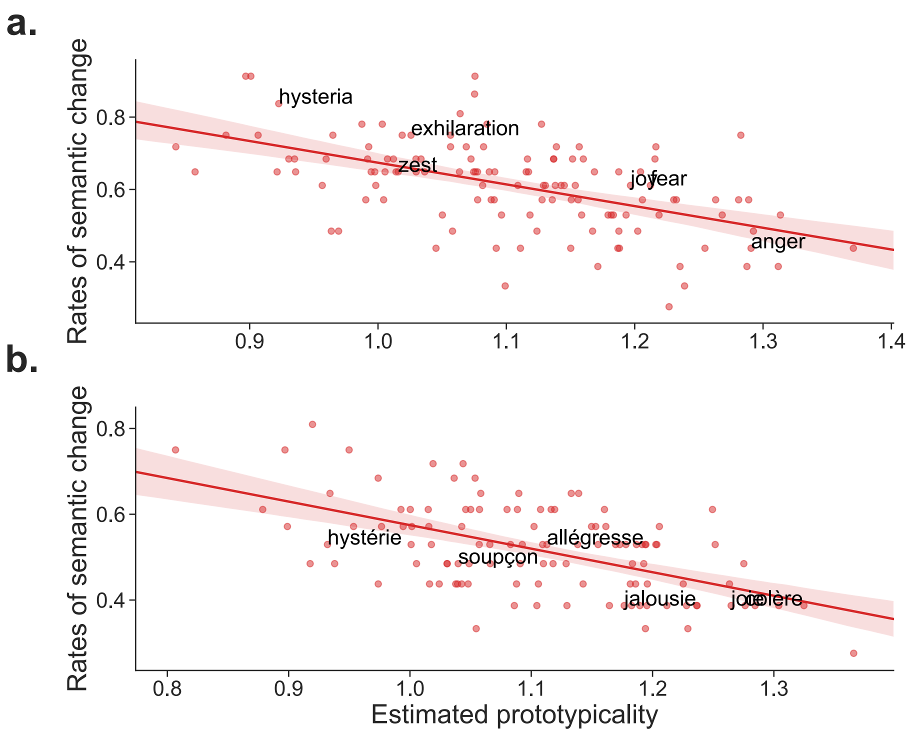

Figure D.11: Scatter plots showing the negative correlations between emotion prototypicality and rates of semantic change between the 1890s and 1990s, in a) English and b) French. Each dot corresponds to an emotion term (with a sample set of words annotated), and each band shows a 95% confidence interval for regressions between prototypicality and rates of semantic change.

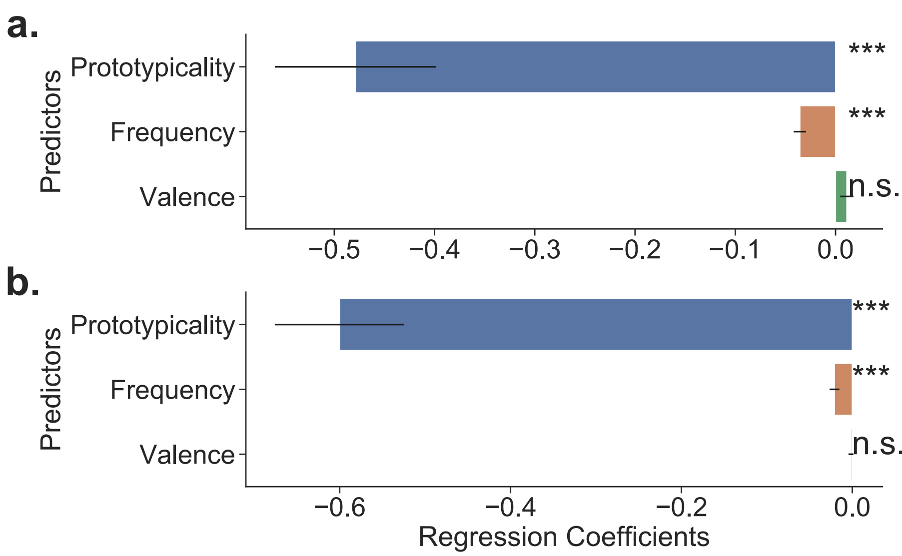

Figure D.12: Predictor coefficients from multiple regressions on rates of semantic change. Error bars show standard error, and "n.s.", "*", "**", "***" denote no significance at $p < 0.05$, and $p < 0.05$, $p < 0.01$, $p < 0.001$ respectively. a) shows results for English, and b) shows results for French.

Appendix D.2. Other factors of semantic change

In the main text, we tested prototypicality as a predictor of semantic stability alongside frequency. Here we examine the role of prototypicality in predicting semantic stability by controlling for three additional predictors: 1) the degree of polysemy of a word, 2) superordinate-subordinate relations between emotion words, and 3) the age of acquisition (AoA) of a word. Firstly, similar to frequency, we control for degree of polysemy because it is a general predictor of semantic change which has been found to negatively correlate with stability in meaning [20, 43]. Secondly, since one function of prototypical emotion words is that they can help define more complex emotion words [11] and this anchoring function may provide relative semantic stability,[4] we examine superordinate-

[4]For example, suppose *joy* is part of the definition of *exhilaration*; if the meaning of *joy* changed, the meaning of *exhilaration* will necessarily change as well, but not vice versa.

subordinate relations (i.e., the hierarchy in the semantic category of emotion) as a potential confounding variable influencing both prototypicality and semantic change. Thirdly, since prototypical emotion words are relatively well-defined and would be easy to learn, and that AoA is a known predictor of stability in lexical change [46, 19], we control for AoA as a potential mediator between prototypicality and semantic change.

Following [43], we operationalized the degree of polysemy of a word as the number of senses the word had at the starting time $t_1 = 1890$ according to the HTE [45]. To operationalize superordinate-subordinate relations, we used WordNet [47] provided by NLTK [48]. Specifically, we constructed a directed graph based on hypernym-hyponym relations, where the root is the sense for *feeling*, and the other nodes correspond to the most frequent sense of an emotion word (see Figure D.13 for illustration). Then, we quantified a word's degree of subordination as its depth in the graph. For example, the word *thrill* has a depth of 4, while *joy* has a depth of 2. Furthermore, to match the historical period of our analyses, we used objective, test-based measurements of AoA originally published by [49]. A digitized version of this data was obtained from [50], where each entry contains a word form, its meaning, and the age at which it was acquired. We computed the AoA of a word by taking the average over all entries in which it appears. Due to the lack of analogous French historical data, we only focused on English emotion words. We assumed these hypernym-hyponym relations and AoA are stable over the past century. After intersecting WordNet and AoA with our historical resources described in the main text, we had a total of 109 English emotion words.

We analyzed these factors using our materials and methods described in *Section 3.1* and *Section 3.2* of the main text. Specifically, we computed the rates of change for every emotion concept x, $rate(x, 1890, 1990)$. Then, we performed multiple regression using the following formula:

$$rate(x, 1890, 1990) \sim p(x|c = emotion) + freq(x) + poly(x) + \tag{D.1}$$

$$val(x) + depth(x) + AoA(x) \tag{D.2}$$

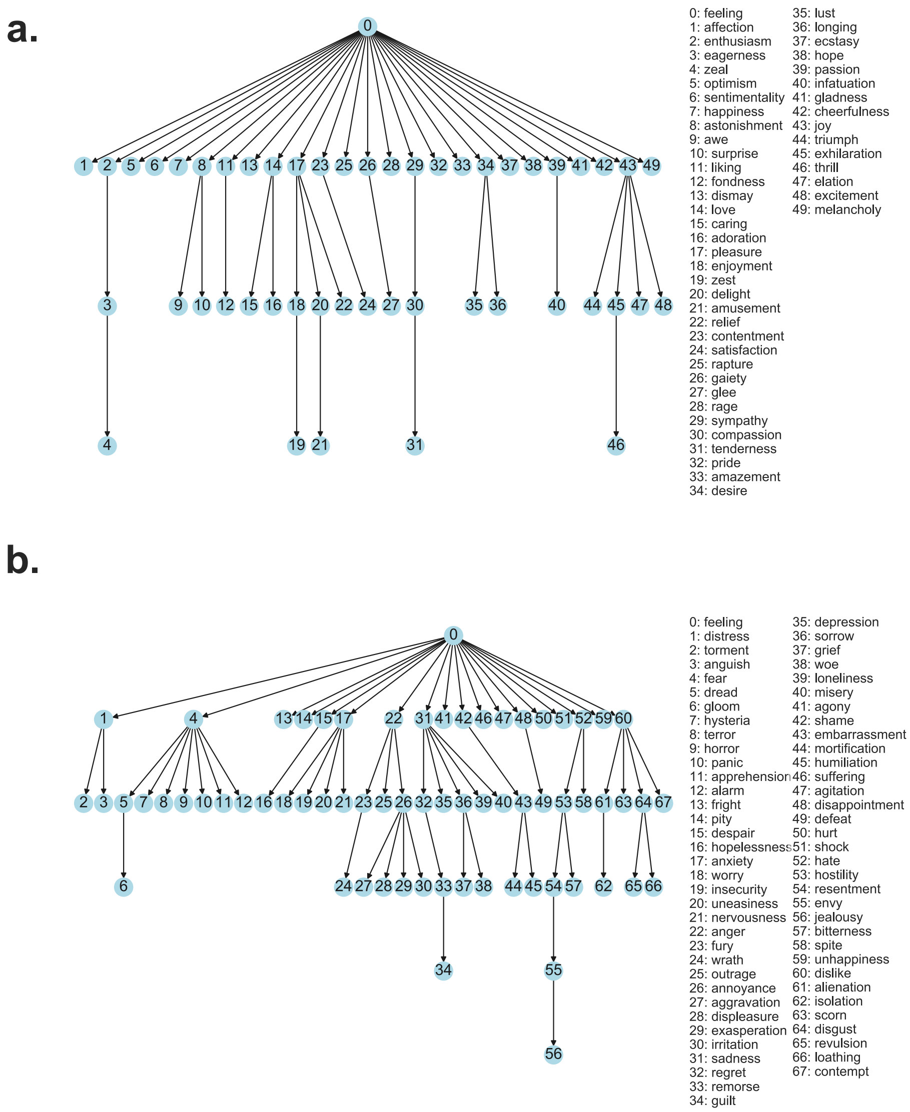

Figure D.13: WordNet hierarchy of hypernyms and hyponyms for a) positive emotion words and b) negative emotion words. Valence is determined by our data described in *Section 3.1*.

where $poly(x)$ is the degree of polysemy of the word operationalized by number of senses, $depth(x)$ is the depth of the concept in the hypernym-hyponym graph we constructed, and $AoA(x)$ is the age at which the word was acquired.

Figure D.14 shows the multiple regression results suggesting the dominant roles of prototypicality and frequency. The adjusted R^2 of the model is 0.697, with $p < 0.001$, $n = 109$; mean regression coefficients for prototypicality ($\beta = -0.4704$, $p < 0.001$) and frequency ($\beta = -0.0460$, $p < 0.001$) are significant, but for valence ($\beta = 0.0049$, $p = 0.267$), number of senses ($\beta = 0.0029$, $p = 0.075$), depth ($\beta = -0.0057$, $p = 0.408$), and AoA ($\beta = 0.0036$, $p = 0.093$) it is insignificant. We observe that prototypicality still has a significant, negative effect as predicted by our hypothesis. We also observe that we can reproduce the finding by [20] for frequency.

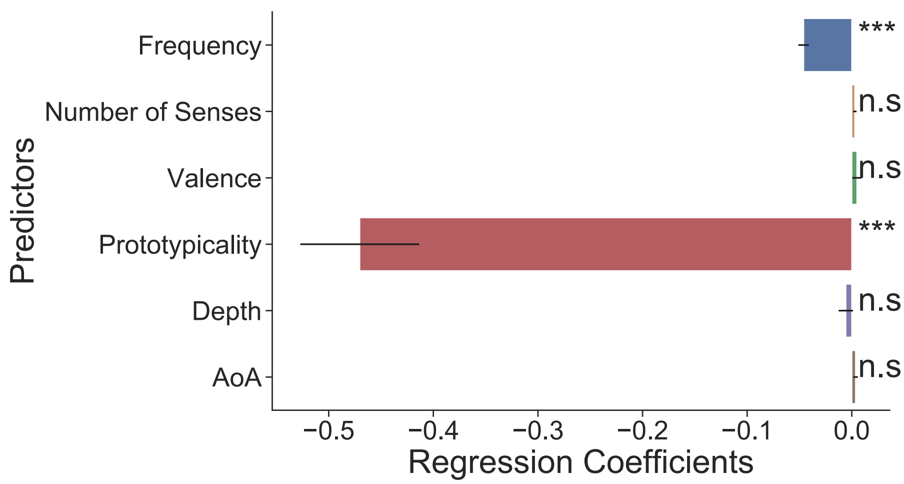

Figure D.14: Predictor coefficients from multiple regressions on rates of semantic change. Error bars follow the same layout as Figure D.12. Prototypicality is estimated using Equation 2 in the main text.

Appendix D.3. Human judgements of prototypicality

We repeated the analysis in the previous section by replacing estimated prototypicality with empirical prototypicality ratings. We computed the rates of change for every emotion concept x, $rate(x, 1980, 1990)$. Then, we performed a multiple regression using the following formula:

$$rate(x, 1980, 1990) \sim proto(x) + freq(x) + poly(x) + val(x) + depth(x) \quad \text{(D.3)}$$

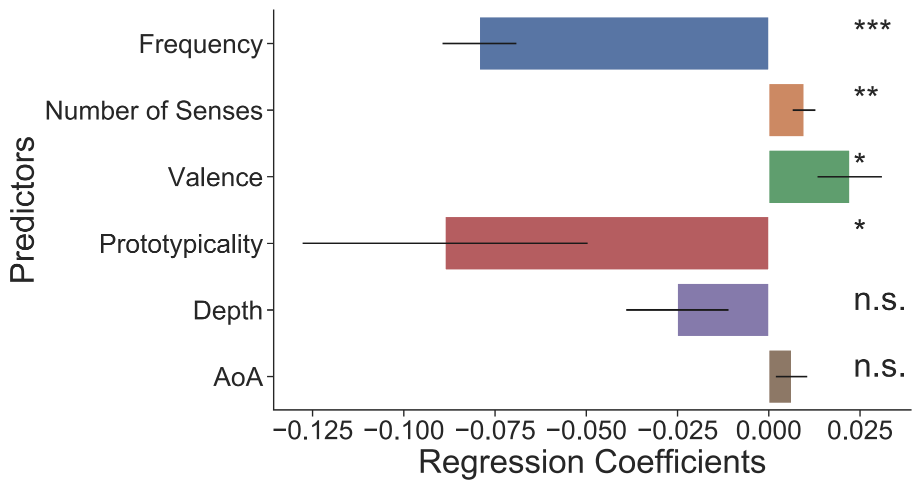

Figure D.15: Predictor coefficients from multiple regressions on rates of semantic change. Error bars follow the same layout as Figure D.12. Prototypicality is based on human ratings.

where $proto(x)$ is the prototypicality rating of x obtained from [6].

Figure D.15 shows the multiple regression results. The adjusted R^2 of the model is 0.524, with $p < 0.001$, $n = 109$; mean regression coefficients for prototypicality ($\beta = -0.0887$, $p = 0.025$), frequency ($\beta = -0.0793$, $p < 0.001$), number of senses ($\beta = 0.0097$, $p = 0.002$), and valence ($\beta = 0.0222$, $p = 0.014$) are significant, but for depth ($\beta = -0.0250$, $p = 0.077$) and AoA ($\beta = 0.0062$, $p = 0.150$) it is insignificant. We observe that empirical prototypicality also has a significant, negative effect, albeit the effect size is smaller than previously.

References

[1] S. E. Kitayama, H. R. E. Markus, Emotion and culture: Empirical studies of mutual influence, American Psychological Association, 1994.

[2] A. Wierzbicka, Emotions across Languages and Cultures: Diversity and Universals, Cambridge University Press, 1999.

[3] J. C. Jackson, J. Watts, T. R. Henry, J.-M. List, R. Forkel, P. J. Mucha, S. J. Greenhill, R. D. Gray, K. A. Lindquist, Emotion semantics show both

cultural variation and universal structure, Science 366 (6472) (2019) 1517–1522.

[4] D. Keltner, J. Haidt, Approaching awe, a moral, spiritual, and aesthetic emotion, Emotion and Cognition 17 (2) (2003) 297–314.

[5] B. Fehr, J. A. Russell, Concept of emotion viewed from a prototype perspective, Journal of Experimental Psychology: General 113 (3) (1984) 464.

[6] P. Shaver, J. Schwartz, D. Kirson, C. O'connor, Emotion knowledge: Further exploration of a prototype approach., Journal of Personality and Social Psychology 52 (6) (1987) 1061.

[7] L. Cosmides, J. Tooby, Evolutionary psychology and the emotions, in: M. Lewis, H.-J. J. M (Eds.), Handbook of Emotions 2nd Edition, Guilford, NY, 2000, pp. 91–115.

[8] L. F. Barrett, Are emotions natural kinds?, Perspectives on Psychological Science 1 (1) (2006) 28–58.

[9] J. A. Russell, L. F. Barrett, Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant., Journal of Personality and Social Psychology 76 (5) (1999) 805.

[10] P. Ekman, D. Cordaro, What is meant by calling emotions basic, Emotion Review 3 (4) (2011) 364–370.

[11] P. N. Johnson-Laird, K. Oatley, The language of emotions: An analysis of a semantic field, Cognition and Emotion 3 (2) (1989) 81–123.

[12] E. Rosch, Cognitive representations of semantic categories., Journal of Experimental Psychology: General 104 (3) (1975) 192.

[13] J. S. Boster, Natural sources of internal category structure: Typicality, familiarity, and similarity of birds, Memory & Cognition 16 (3) (1988) 258–270.

[14] P. Garrard, M. A. Lambon Ralph, J. R. Hodges, K. Patterson, Prototypicality, distinctiveness, and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts, Cognitive Neuropsychology 18 (2) (2001) 125–174.

[15] E. Lieberman, J.-B. Michel, J. Jackson, T. Tang, M. A. Nowak, Quantifying the evolutionary dynamics of language, Nature 449 (7163) (2007) 713–716.

[16] M. Pagel, Q. D. Atkinson, A. Meade, Frequency of word-use predicts rates of lexical evolution throughout Indo-European history, Nature 449 (7163) (2007) 717–720.

[17] M. Pagel, Q. D. Atkinson, A. S. Calude, A. Meade, Ultraconserved words point to deep language ancestry across Eurasia, Proceedings of the National Academy of Sciences 110 (21) (2013) 8471–8476.

[18] S. Vejdemo, T. Hörberg, Semantic factors predict the rate of lexical replacement of content words, PLOS ONE 11 (1) (2016) e0147924.

[19] P. Monaghan, S. G. Roberts, Cognitive influences in language evolution: Psycholinguistic predictors of loan word borrowing, Cognition 186 (2019) 147–158.

[20] W. L. Hamilton, J. Leskovec, D. Jurafsky, Diachronic word embeddings reveal statistical laws of semantic change, in: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 2016.

[21] H. Dubossarsky, D. Weinshall, E. Grossman, Outta control: Laws of semantic change and inherent biases in word representation models, in: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017, pp. 1136–1145.

[22] R. Boyd, P. J. Richerson, Culture and the evolutionary process, University of Chicago press, 1988.

[23] J. Bybee, et al., Frequency of use and the organization of language, Oxford University Press on Demand, 2007.

[24] D. Geeraerts, Diachronic Prototype Semantics: A Contribution to Historical Lexicology, Oxford University Press, 1997.

[25] C. Ramiro, M. Srinivasan, B. C. Malt, Y. Xu, Algorithms in the historical emergence of word senses, Proceedings of the National Academy of Sciences 115 (10) (2018) 2323–2328.

[26] H. Dubossarsky, Y. Tsvetkov, C. Dyer, E. Grossman, A bottom up approach to category mapping and meaning change., in: NetWordS, 2015, pp. 66–70.

[27] T. Mikolov, K. Chen, G. S. Corrado, J. Dean, Efficient estimation of word representations in vector space, in: Workshop Papers in International Conference on Machine Learning, 2013.

[28] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, in: Advances in Neural Information Processing Systems, 2013, pp. 3111–3119.

[29] A. Caliskan, J. J. Bryson, A. Narayanan, Semantics derived automatically from language corpora contain human-like biases, Science 356 (6334) (2017) 183–186.

[30] V. Kulkarni, R. Al-Rfou, B. Perozzi, S. Skiena, Statistically significant detection of linguistic change, in: Proceedings of the 24th International Conference on World Wide Web, 2015.

[31] J. Xie, R. Ferreira Pinto Jr., G. Hirst, Y. Xu, Text-based inference of moral sentiment change, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, 2019.

[32] Y. Li, T. Engelthaler, C. S. Siew, T. T. Hills, The macroscope: A tool for examining the historical structure of language, Behavior Research Methods 51 (4) (2019) 1864–1877.

[33] S. Buechel, U. Hahn, Word emotion induction for multiple languages as a deep multi-task learning problem, in: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2018, pp. 1907–1918.

[34] S. M. Mohammad, Sentiment analysis: Detecting valence, emotions, and other affectual states from text, in: Emotion Measurement, 2016, pp. 201–237.

[35] R. A. Calvo, S. Mac Kim, Emotions in text: dimensional and categorical models, Computational Intelligence 29 (3) (2013) 527–543.

[36] Y. Xu, C. Kemp, A computational evaluation of two laws of semantic change., in: Proceedings of the 37th Annual Meeting of the Cognitive Science Society, 2015.

[37] P. Niedenthal, C. Auxiette, A. Nugier, N. Dalle, P. Bonin, M. Fayol, A prototype analysis of the french category "émotion", Cognition and Emotion 18 (3) (2004) 289–312.

[38] S. K. Reed, Pattern recognition and categorization, Cognitive Psychology 3 (3) (1972) 382–407.

[39] F. G. Ashby, L. A. Alfonso-Reese, Categorization as probability density estimation, Journal of Mathematical Psychology 39 (2) (1995) 216–233.

[40] S. Seabold, J. Perktold, statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference, 2010.

[41] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometrics and Intelligent Laboratory Systems 2 (1-3) (1987) 37–52.

[42] E. Rosch, Principles of categorization, in: E. Rosch, B. B. Lloyd (Eds.), Cognition and Categorization, Lawrence Erlbaum Associates, Hillsdale, NJ, 1978, pp. 27–48.

[43] Y. Luo, Y. Xu, Stability in the temporal dynamics of word meanings., in: Proceedings of the 40th Annual Meeting of the Cognitive Science Society, 2018.

[44] A. Xu, C. Ramiro, Y. Xu, A predictability-distinctiveness trade-off in the historical emergence of word forms., in: Proceedings of the 41st Annual Meeting of the Cognitive Science Society, 2019.

[45] C. Kay, J. Roberts, M. Samuels, I. Wotherspoon, The Historical Thesaurus of English, version 4.21 (2017).
URL http://historicalthesaurus.arts.gla.ac.uk/

[46] P. Monaghan, Age of acquisition predicts rate of lexical evolution, Cognition 133 (3) (2014) 530–534.

[47] C. Fellbaum, WordNet: An Electronic Lexical Database, MIT Press, Cambridge, MA, 1998.

[48] E. Loper, S. Bird, Nltk: The natural language toolkit, in: In Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics. Philadelphia: Association for Computational Linguistics, 2002.

[49] E. Dale, J. O'rourke, The Living Word Vocabulary, the Words We Know: A National Vocabulary Inventory., Chicago: World Book, 1981.

[50] M. Brysbaert, A. Biemiller, Test-based age-of-acquisition norms for 44 thousand english word meanings, Behavior research methods 49 (4) (2017) 1520–1523.