A Computational Evaluation of Two Laws of Semantic Change

Yang Xu (yang_xu_ch@berkeley.edu)
Department of Linguistics
University of California, Berkeley

Charles Kemp (ckemp@cmu.edu)
Department of Psychology
Carnegie Mellon University

Abstract

For more than a century scholars have proposed laws of semantic change that characterize how words change in meaning over time. Two such laws are the law of differentiation, which proposes that near-synonyms tend to differentiate in meaning over time, and the law of parallel change, which proposes that related words tend to undergo parallel changes in meaning. Researchers have identified a handful of changes that are consistent with each proposed law, but there are no systematic evaluations that assess the validity and generality of these competing laws. Here we evaluate these laws by using a large corpus to assess how thousands of related words changed in meaning over the twentieth century. Our analyses show that the law of parallel change applies more broadly than the law of differentiation, and thereby illustrate how large-scale computational analyses can place laws of semantic change on a more secure footing.

Keywords: semantic change; law of differentiation; law of parallel change; computational semantics

Introduction

The sounds, structures and meanings of languages are constantly changing. Shakespeare would have sounded different from a modern English speaker, and his plays demonstrate that both grammar and word meanings have changed since the turn of the seventeenth century. Scholars have formulated laws that characterize the nature of these changes, and the literature contains laws of sound change (Labov, 2010), grammatical or syntactic change (Hopper & Traugott, 2003; Lieberman, Michel, Jackson, Tang, & Nowak, 2007), and semantic change (Bréal, 1897; Sweetser, 1991; Traugott & Dasher, 2002). Semantic change, however, is the area of historical linguistics that is "least well understood" (p 197 in Crowley and Bowern, 2010), and proposed laws of semantic change have not achieved the same secure status as laws of sound change.

Here we focus on two laws of semantic change that are prominent but incompatible. The law of differentiation (Bréal, 1897; Sturtevant, 1917) proposes that near-synonyms tend to diverge in meaning over time. Figures 1a and 1b show an example suggesting that "fragile" and "frail" were neighbors in semantic space during the 1890s but had moved apart by the 1990s. Many researchers have proposed that it is inefficient for languages to contain multiple forms with similar meanings (Bolinger, 1977), and selective pressures toward increased efficiency provide a plausible reason why differentiation may occur. The second law is the law of parallel change, which proposes that words with related meanings tend to change in similar ways over time (Stern,

1921; Lehrer, 1985). Figures 1c and 1d suggest that the meanings of "imminent" and "impending" changed in similar ways between the 1890s and the 1990s. One plausible explanation for parallel change is that language users tend to preserve associations between closely related words, which means that a word that changes in meaning is likely to "drag along" other related words (p 286 in Lehrer, 1985).

The laws of differentiation and parallel change make opposite predictions about how near-synonyms change over time. Which prediction is correct will vary from case to case, and linguists have identified examples of change that support the law of differentiation (Bréal, 1897; Sturtevant, 1917) and other examples that support the law of parallel change (Stern, 1921; Lehrer, 1985). These sets of examples, however, are typically very small, and are unable to reveal whether each law applies in general or only in rare cases. To assess the generality and importance of these competing laws, we present a large-scale computational analysis that explores how thousands of pairs of related words change in meaning over time. One group of researchers has previously argued that laws of semantic change are statistical tendencies that need to be evaluated using large-scale statistical analyses (Williams, 1976; Ullmann, 1943). A second group of researchers has explored computational methods for detecting semantic change in large corpora (Sagi, Kaufmann, & Clark, 2011; Gulordava & Baroni, 2011). We bring these ideas together by showing how large-scale computational analyses can be used to evaluate proposed laws of semantic change.

Results

Our analysis explores how nouns, verbs, and adjectives changed in meaning over the eleven decades between 1890 and 1999. We use the Google Million corpus (Michel et al., 2010)[1], which contains around 650 million words for each year in the period that we consider. For all of our analyses we binned the data into decades. We used a distribuitional approach to meaning (Firth, 1957), and captured the meaning of a word during a given decade by a *meaning vector* that reflects the contexts in which it appeared. The *Materials and Methods* section describes how we converted these raw vectors into normalized meaning vectors, or probability distributions that sum to 1. To measure the similarity between two meaning vectors, we used the Jensen-Shannon (JS) diver-

---

Differentiation

Parallel change

a 1890s

brittle · disabled · precarious
stringy · diminutive · tiny · feeble · senile
flimsy · delicate · vulnerable
sturdy · slender · healthy · clumsy · wretched
robust · infirm
demented · needy

△ fragile
▽ frail

1990s

brittle · tiny · precarious
healthy · delicate · robust
stringy · flimsy · diminutive · clumsy · feeble
sturdy · slender · wretched
disabled · senile · vulnerable
infirm · demented · needy

Semantic space

b 1890s   1990s

__bark
__construct
__mortal
__china
__boat
__mortality
__humanity
__flower
__shell
__build
__frame
concerns __
__earthly
bones __
__hold
__shoulders
__elderly
__bones
__aged
__alliance
__elder
__cease
__older

□ fragile
■ frail

0   0.2   0.4   0   0.2   0.4
Relative frequency

c 1890s

rife · raging · ensuing
inevitable · eventual · incipient
menacing · sudden · gradual
serious · unexpected · tardy
appalling · disastrous
terrible · ignominious
fearful · dreadful

△ imminent
▽ impending

1990s

eventual · ensuing
sudden · gradual
serious · incipient
menacing
raging · appalling · unexpected
terrible · ignominious
disastrous · rife
fearful · dreadful · inevitable
tardy

Semantic space

d 1890s   1990s

__danger
danger __
crisis __
__risk
dangers __
ruin __
destruction __
fate __
revolution __
conflict __
__ecological
invasion __
__ruin
__death
__invasion
__collapse
election __
__arrival
__departure
__coming
__divorce

□ imminent
■ impending

0 0.1 0.2 0.3   0 0.1 0.2 0.3
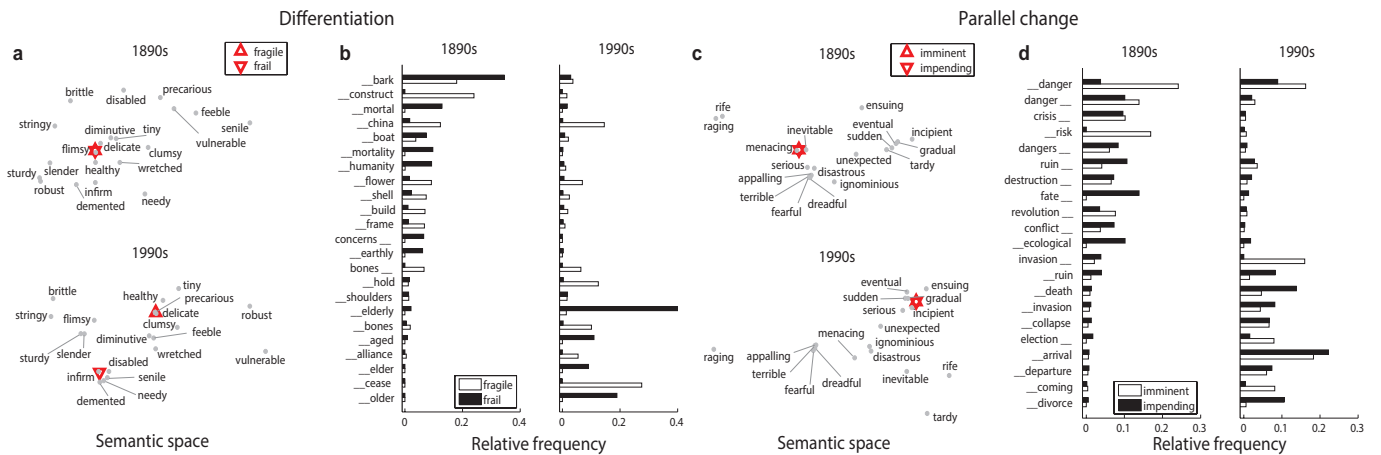Relative frequency

Figure 1: Word pairs that illustrate differentiation and parallel change. (a) Semantic neighbors of "fragile" and "frail" in the 1890s and the 1990s. The two words begin as nearest neighbors but move apart over a century and end up closest to "delicate" and "infirm" respectively. The semantic spaces shown were created from the Google Ngram corpus (Michel et al., 2010) using the t-SNE algorithm for dimensionality reduction (Maaten & Hinton, 2008). (b) Distributions showing the contexts that are linked most strongly with "fragile" and "frail." Over the century, "frail" alone becomes strongly linked with "elderly" and "older." (c) Semantic neighbors of "imminent" and "impending" in the 1890s and the 1990s. Both words start out closest to "menacing" but end up close to "gradual" and "incipient." (d) Context distributions for "imminent" and "impending." Over the century, both words become linked more strongly with words like "arrival" and "departure" that do not convey a sense of danger.

gence, which is a standard measure of the distance between two distributions. If two words appeared in the same contexts during a given decade, then the JS divergence between their meaning vectors is small.

Our first analysis focused on sets of English synonyms collected from two historical resources published in 1896 (Fernald, 1907) and 1920 (Allen, 1920) respectively. We began by asking whether synonym pairs were more likely to move apart in semantic space than control pairs including words that were not necessarily related in meaning. The law of differentiation predicts that synonyms should tend to diverge more than control pairs, but the law of parallel change predicts that synonyms should tend to stay closer than the controls. We tested these predictions using synonym pairs where both words changed more than the population average over the eleven decades. The degree of semantic change for a given word was computed by finding the word's 100 nearest neighbors in the 1890s and again in the 1990s, and calculating the overlap between these sets. For each synonym pair we chose a control pair that satisfied two criteria. First, the JS divergence between the two control words in the 1890s must be smaller than the JS divergence between the synonyms for the same decade. Second, the total amount of semantic change over the eleven decades must be smaller for the control words than the synonyms. Because the words in each control pair are initially nearby in semantic space, some of the control pairs are semantically related. For example, the controls for "imminent" and "impending" in Figure 1 are "instructive" and "interesting," and the controls for "fragile" and "frail" are "optimistic" and "pessimistic." The controls, how-

ever, also include pairs such as "lonely" and "western" that are not semantically related but that happen to have similar context distributions. Our policy for choosing controls ensures that the control pairs start closer to each other and move less than the synonyms in semantic space. All other things being equal, we should therefore expect the control pairs to stay closer together than the synonym pairs.

Figure 2 shows, however, that synonym pairs in both historical sets tend to stay closer than the control pairs. The plots are based on pairs of synonyms that are divided into nouns, adjectives, and verbs. In all plots, our conservative policy for choosing controls ensures that the average distance between control pairs is initially smaller than the average distance between synonyms, but by the 1990s the control pairs are further apart on average than the synonyms. To evaluate the statistical significance of this result, we compared the number of cases where the synonyms ended up closer with the number of cases where the control pairs ended up closer. Figure 2b and Figure 2f show that binomial tests from all word groups yielded significant results in the Fernald source (nouns $p < 0.0003, n = 222$, adjectives $p < 0.001, n = 142$ and verbs $p < 0.02, n = 41$) and in the Allen source (nouns $p < 0.04, n = 790$ and verbs $p < 0.05, n = 304$) except for the adjective group ($p = 0.31; n = 344$). Our results therefore provide evidence against the law of differentiation and in favor of the law of parallel change.

The law of parallel change should also apply to antonyms, and we repeated our first analysis using antonym pairs drawn from the same historical sources. The results in Figures 2c-d and 2g-h provide additional support for the law of parallel
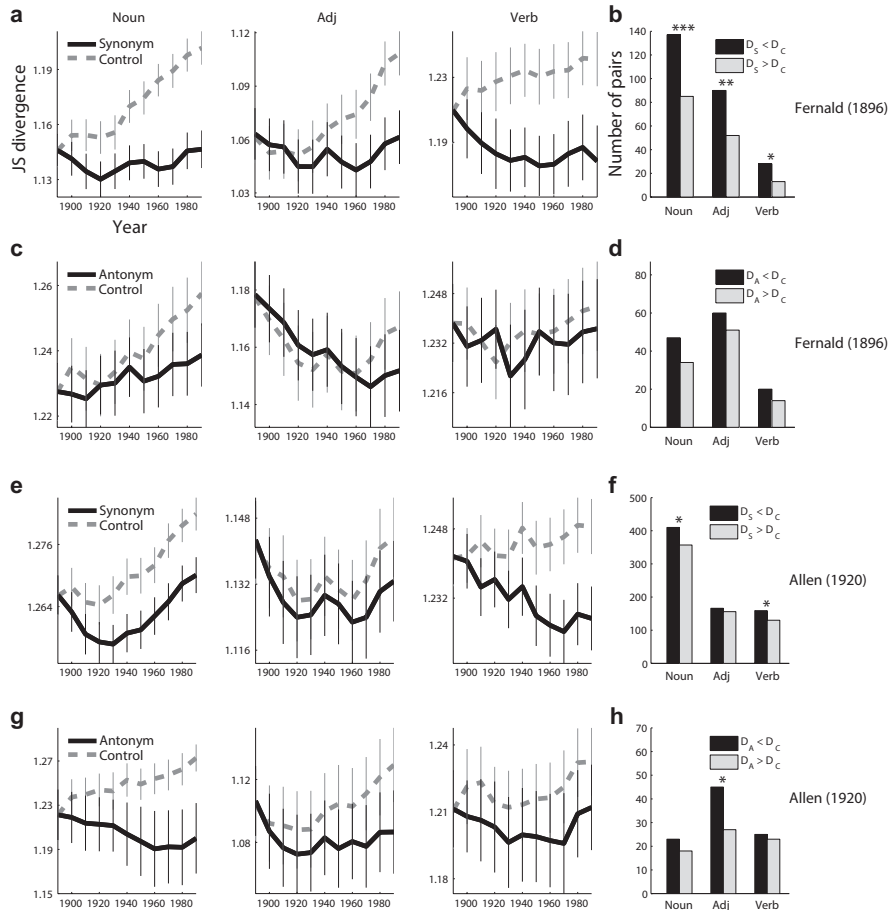
Figure 2: Forward change in synonym pairs, antonym pairs and control pairs from the 1890s to the 1990s. (a) Pairwise JS divergences between control pairs and synonymous nouns, adjectives, and verbs from a 1896 publication (Fernald, 1907). (b) Counts showing the number of synonym pairs that end up closer ($D_S < D_C$) or further apart ($D_S > D_C$) than their control pairs by the 1990s. (c) JS divergences between control pairs and antonym pairs. (d) Counts showing whether antonym pairs tend to end up closer ($D_A < D_C$) or further apart ($D_A > D_C$) than their control pairs. (e), (f), (g), (h) Analysis of synonym and antonym pairs from a 1920 publication (Allen, 1920). All error bars indicate the standard error of the mean, and "*", "**" and "***" indicate statistical significance at $p < 0.05$, 0.005 and 0.0005 respectively.

change by suggesting that antonym pairs are likely to stay closer in semantic space than control pairs. This effect is less robust for antonyms than synonyms, possibly because the number of antonym pairs is smaller. Overall, however, Figure 2 shows that there are more cases where antonym pairs stay closer together than controls in both historical sources (nouns $p = 0.09, n = 81$, adjectives $p = 0.22, n = 111$ and verbs $p = 0.20, n = 34$ in the Fernald source; nouns $p = 0.27, n = 42$, adjectives $p < 0.03, n = 72$ and verbs $p = 0.44, n = 48$ in the Allen source).

Although our analyses so far support the law of parallel change, it is possible that our results are shaped in part by aspects of the Google million corpus that are unrelated to semantic change. In particular, the 1990s data were sampled from a more diverse collection of sources than were the 1890s data, and this difference in diversity could potentially affect analyses that explore how word meanings change over time. To rule out this possibility, we ran an analysis that reversed the direction of time and explored whether pairs that

were related in the 1990s were likely to stay nearby in semantic space as time was rolled backwards. Although any causal forces that keep related pairs together operate forward in time, a backwards analysis can still provide evidence for the existence of these forces. For example, if two words are synonyms in the 1990s, the law of parallel change predicts that the words will have arrived at their current locations in semantic space by following parallel trajectories, which implies that the words should be nearby in semantic space during the 1890s.

To test this prediction, we collected sets of synonyms from two modern resources developed during the 1990s: *Word-Net* (Miller, 1995) and the *Moby Thesaurus* (Ward, 2002). Our analyses followed the same general procedure as our previous analyses, and the results are shown in Figure 3. Because we are interested in changes that occur as time is rolled backwards, the time axis has been reversed in Figures 3a, 3c and 3e. In all cases, synonyms were more likely than control pairs to stay together as time is rolled backwards, and bi-
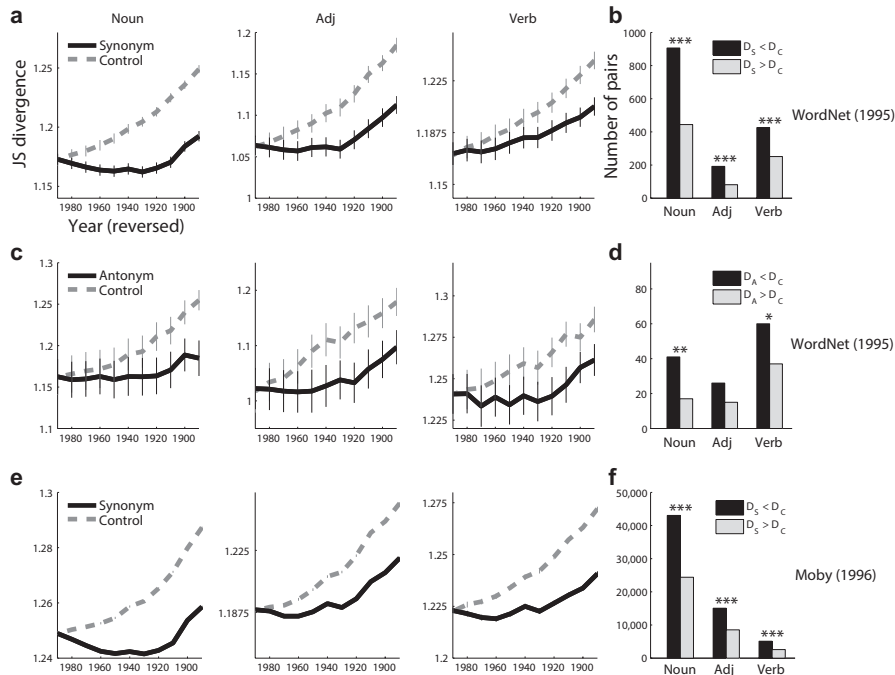
Figure 3: Backward change in synonym pairs, antonym pairs and control pairs from the 1990s to the 1890s. (a) Pairwise JS divergences between control pairs and synonym pairs from WordNet (Miller, 1995) published in 1995. (b) Counts showing the number of synonym pairs that end up closer ($D_S < D_C$) or further apart ($D_S > D_C$) than their control pairs by the 1890s. (c) JS divergences between control pairs and antonym pairs. (d) Counts showing whether antonym pairs tend to end up closer ($D_A < D_C$) or further apart ($D_A > D_C$) than their control pairs. (e), (f) Analysis of synonym pairs drawn from the Moby thesaurus (Ward, 2002) published in 1996.

nomial tests from all word groups yielded highly significant results ($p < 10^{-11}$ across 1350 noun, 273 adjective and 677 verb pairs in WordNet; $p < 10^{-11}$ across 67482 noun, 23591 adjective and 7698 verb pairs in the Moby set). The Moby thesaurus does not include antonyms, but we used antonym pairs from WordNet to explore whether parallel change applies to antonyms as time is rolled backwards. Figures 3c and 3d show that antonym pairs tend to stay closer than control pairs (nouns: $p < 0.002, n = 58$; adjectives: $p = 0.06, n = 41$; verbs: $p < 0.02, n = 97$). Regardless of whether the 1990s are the start point or the end point of our analyses, we therefore find strong support for the law of parallel change, which suggests that the results in Figures 2 and 3 are not artifacts of the corpus that we used.

Discussion

The analyses described in this paper provide consistent evidence in favor of the law of parallel change and against the law of differentiation. Our results, however, do not provide direct evidence about the causal mechanisms that produce parallel change. One theory proposes that parallel change results from cognitive forces such as analogy that act to preserve patterns of relationships between words (Anttila, 1977). On this view, words which change in meaning directly affect the meaning of words to which they are mentally associated (Lehrer, 1985; Kroesch, 1926). Another theory proposes that words do not directly affect other related words, but that groups of related words are affected in similar ways by exter-

nal forces (Stern, 1931), including non-linguistic forces such as social and technological change. Both of these theories seem plausible and future studies are needed to choose between them.

Our results do not imply that synonyms never differentiate in meaning. Differentiation may be especially likely to occur when languages that contain many near-synonyms come into contact. For example, the English words *hound* and *dog* moved apart in meaning after dog was imported from Scandinavian (Bréal, 1897). Our analysis focused on English from the 1890s on, and language contact shaped English to a much greater extent during earlier periods including the centuries following the Norman Conquest in 1066. We cannot draw any conclusions about the relative frequencies of differentiation and parallel change during periods of sustained language contact, but our results suggest that parallel change is more common than differentiation outside of these periods.

The literature contains many proposals about laws of semantic change, but there have been few comprehensive attempts to evaluate these laws. Our work demonstrates how computational analyses of large corpora can be used to evaluate proposed laws of semantic change. As yet, theories of semantic change are still relatively undeveloped compared to theories that explain how the sound inventories of languages change over time. Large scale computational approaches, however, may ultimately lead to laws of semantic change that are just as well supported as laws of sound change.

Materials and Methods

Target and context words. We chose a set of *target words* and an independent set of *context words*. The target words were classified as nouns, verbs or adjectives. Because the Google Million does not include part-of-speech tags, we selected these target words from the Corpus of Historical American English (COHA) (Davies, 2011). COHA contains approximately 400 million words that appeared at least 3 times between 1810 and 2009. We used a case-insensitive list of these words that included frequencies and part-of-speech (POS) tags.[2] If a word was listed with multiple POS tags, we stripped all but the most probable form of the word, which we defined as the form with maximal median frequency over the eleven decades in our analysis. In addition, words consisting of a single letter, words that appeared less than once on average per year, and words that co-occurred with fewer than 1% (50) of the context words were removed. The set that remains included 9,886 nouns, 3,431 adjectives, and 5,022 verbs, which makes 18,339 target words in total. The context words were 5000 words that appeared frequently during both 1890 and 1999. We selected these words by collecting the 3868 most frequent words from 1890 and the 3868 most frequent words from 1999. These two sets of words included 2736 words in common, and combining the two sets produced a set of 5000 unique context words. Because we consider analyses that run both forwards and backwards in time, the context words were chosen to be equally representative of the years at the beginning and end of the 11 decades.

Meaning vectors. We captured the meaning of a word during a given decade by a 10,000 element *meaning vector*. The raw vector for each word in each decade specifies the number of times that the word appeared immediately to the left and right of each of 5000 context words. The resulting counts are organized into separate matrices for target nouns (9,886 by 10,000), adjectives (3,431 by 10,000) and verbs (5,022 by 10,000) where each row in the matrix corresponds to a target word, and each column corresponds to a context word that either precedes or follows a target word slot. All words were converted to lower case before computing these counts. The meaning of a target word w_i during a given decade is represented as a vector $\mathbf{v}_i^t$, where t denotes the time (i.e. decade) in question. Our analyses use meaning vectors that reflect the extent to which each word co-occurs with our set of 5000 context words. The meaning vectors for nouns, adjectives and verbs during any given decade are created by starting with the corresponding matrix of co-occurrence counts for that decade, then normalizing the columns so that they sum to one. Normalizing in this way ensures that changes in the frequencies of the context words will not affect the meaning vectors for the target words. We then take the normalized matrix and normalize once more so that the rows sum to one. The rows of this doubly-normalized matrix are the meaning vectors used in our analyses, and each vector can be viewed as a probability distribution over contexts.

Measuring semantic distances between words. For any two target words w_i and w_j, let $D^t(w_i, w_j)$ denote the semantic distance between these words at time t. Intuitively, $D^t(w_i, w_j)$ should be small to the extent that the meaning vectors $\mathbf{v}_i^t$ and $\mathbf{v}_j^t$ are similar. We measure semantic distance using the Jensen–Shannon (JS) divergence:

$$D^t(w_i, w_j) = \frac{1}{2}\left(KL(\mathbf{v}_i^t||\mathbf{m}^t) + KL(\mathbf{v}_j^t||\mathbf{m}^t)\right), \quad (1)$$

where $\mathbf{m}^t = \frac{1}{2}(\mathbf{v}_i^t + \mathbf{v}_j^t)$, KL is the Kullback-Leibler divergence

$$KL(\mathbf{v}_i^t||\mathbf{m}^t) = \sum_c v_i^t(c) \cdot \log v_i^t(c) - \sum_c v_i^t(c) \cdot \log m^t(c) \quad (2)$$

and both sums are over all contexts c in the meaning vectors.

Measuring degrees of semantic change. Equation 1 is used to measure the distance between two different words during the same decade. To quantify how much a single word changes in meaning over the eleven decades, we measured the degree to which that word moves around relative to its neighbors in semantic space. Intuitively, a word has changed in meaning if its nearest neighbors in 1890 do not overlap substantially with its nearest neighbors in 1999. We capture this idea by using Equation 1 to compute semantic distances between every pair of target words during the 1890s, and again during the 1990s. We used these distances to identify the nearest 100 neighbors for each target word during the 1890s and again during the 1990s. For each target word, we then compute the proportion of the 1990s neighbors that were also neighbors during the 1890s. The greater the amount of semantic change, the smaller the proportion of shared neighbors. The degree of semantic change is therefore defined as 1 minus the proportion of shared neighbors. We treat nouns, adjectives and verbs separately. For example, when computing the 100 nearest neighbors for a given noun, we consider only neighbors that are nouns.

Sources of synonym and antonym pairs. The synonyms and antonyms used in our analyses were collected from two historical and two modern sources. The first historical source is a book from 1896 called *English Synonyms and Antonyms* (Fernald, 1907). We used the project Gutenberg version of the book.[3] The book is organized around a set of headwords, and synonyms and antonyms are listed for each headword. We created a list of synonym pairs by pairing each headword with each listed synonym, and created a list of antonym pairs similarly. We pruned all pairs that included one or more words that did not appear among our target words, and classified the pairs as nouns, adjectives or verbs based on the part-of-speech tags included in our list of target words. In addition, we pruned pairs where the degree of semantic change between the 1890s and 1990s was below average for both words. These procedures yielded a total of 222 synonym and 81 antonym noun pairs, 142 synonym and 111 antonym adjective pairs, and 41 synonym and 34 antonym verb pairs. The second historical source is a book from 1920

called *Allen's Synonyms and Antonyms* (Allen, 1920). We used the version in the Internet Archive from the University of California Libraries.[4] The book is also organized around a set of headwords with their corresponding synonyms and antonyms. Using similar procedures, we pruned all pairs that did not appear among our target words, and classified remaining pairs as nouns, adjectives or verbs based on the part-of-speech tags from our list of target words. We also pruned pairs where the degree of semantic change between the 1890s and 1990s was below average for both words. These procedures yielded a total of 790 synonym and 42 antonym noun pairs, 344 synonym and 72 antonym adjective pairs, and 304 synonym and 48 antonym verb pairs.

The first modern set was collected from *WordNet* (Miller, 1995) which includes both synonyms and antonyms. In particular, synsets and antonyms for all words in our target sets were extracted (as of March 17, 2012) using the Natural Language Toolkit corpus reader (http://nltk.org/). We pruned all pairs that did not appear in our set of target words and classified the pairs as nouns, adjectives or verbs based on the part-of-speech tags included in our list of target words. In addition, we pruned all pairs where the degree of semantic change was below average for both words. The remaining set includes 1350 synonym and 58 antonym noun pairs, 273 synonym and 41 antonym adjective pairs, and 677 synonym and 91 antonym verb pairs. The second modern set of synonyms was collected from the *Moby Thesaurus* (Ward, 2002).[5] We combined all root words in the thesaurus with their associated words, and stripped all words that did not appear in our set of target words. In addition, we pruned pairs where both words show below-average degree of semantic change. The final collection of synonyms included 67,482, 23,591 and 7,698 pairs of nouns, adjectives and verbs.

Acknowledgments

References

Allen, F. S. (1920). *Allen's synonyms and antonyms*. New York and London: Harper and Brothers Publishers.

Anttila, R. (1977). *Analogy*. Berlin: Mouton de Gruyter.

Bolinger, D. (1977). *Meaning and form*. London: Longman.

Bréal, M. (1897). *Essai de sémantique*. Paris: Hachette.

Crowley, T., & Bowern, C. (2010). *An introduction to historical linguistics*. Oxford: Oxford University Press.

Davies, M. (2011). N-grams and word frequency data from the Corpus of Historical American English (COHA). Downloaded from http://www.ngrams.info on Jan 9, 2012.

Fernald, J. C. (1907). *English synonyms and antonyms*. New York: Funk & Wagnalls company.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. In *Studies in linguistic analysis (special volume of the Philological Society)*. Oxford: Blackwell.

Gulordava, K., & Baroni, M. (2011). A distributional similarity approach to the detection of semantic change in the Google Books Ngram corpus. In *Proceedings of the GEMS 2011 workshop*. Edinburgh, Scotland.

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.

Kroesch, S. (1926). Analogy as a factor in semantic change. *Language*, *2(1)*, 35–45.

Labov, W. (2010). *Principles of Linguistic Change, Volume III, Cognitive and Cultural Factors*. New Jersey: Wiley-Blackwell.

Lehrer, A. (1985). The influence of semantic fields on semantic change. In J. Fisiak (Ed.), *Historical semantics: Historical word formation* (pp. 283–296). Berlin: Mouton de Gruyter.

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

Maaten, L. J. P. van der, & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *JMLR*, *9*, 2579–2605.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Brockman, W., et al. (2010). Quantitative analysis of culture using millions of digitized books. *Science*, *331(6014)*, 176–182.

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, *38(11)*, 39–41.

Sagi, E., Kaufmann, S., & Clark, B. (2011). Tracing semantic change with latent semantic analysis. In *Current methods in historical semantics*. Berlin: Mouton de Gruyter.

Stern, G. (1921). *Swift, swiftly, and their synonyms: A contribution to semantic analysis and theory*. Göteberg: Wettergren & Kerber.

Stern, G. (1931). *Meaning and change of meaning, with special reference to the English language*. Bloomington: Indiana University Press.

Sturtevant, E. H. (1917). *Linguistic change: An introduction to the historical study of language*. Chicago: The University of Chicago Press.

Sweetser, E. (1991). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.

Traugott, E. C., & Dasher, R. B. (2002). *Regularities in semantic change*. Cambridge: Cambridge University Press.

Ullmann, S. (1943). Laws of language and laws of nature. *The Modern Language Review*, *38(4)*, 328–338.

Ward, G. (2002). Moby thesaurus II. *Project Gutenberg Literary Archive Foundation*.

Williams, J. M. (1976). Synaesthetic adjectives: A possible law of semantic change. *Language*, *52(2)*, 461–478.

---

[4] Downloaded from http://archive.org/details/allenssynonymsan00alle on March 15, 2013.

[5] Downloaded from http://www.gutenberg.org/ebooks/3202 on October 20, 2012.