12

ATTENTION ECONOMIES, INFORMATION CROWDING, AND LANGUAGE CHANGE

Thomas T. Hills
James Adelman
Takao Noguchi

Abstract

Language is a communication system that adapts to the cognitive capacities of its users and the social environment in which it used. In this chapter we outline a theory, inspired by the linguistic niche hypothesis, which proposes that language change is influenced by information crowding. We provide a formal description of this theory and how information markets, caused by a growth in the communication of ideas, should influence conceptual complexity in language over time. Using American English and data from multiple language corpora, including over 400 billion words, we test the proposed crowding hypothesis as well as alternative theories of language change, including learner-centered accounts and semantic bleaching. Our results show a consistent rise in concreteness in American English over the last 200 years. This rise is not strictly due to changes in syntax, but occurs within word classes (e.g. nouns), as well as within words of a given length. Moreover, the rise in concreteness is not correlated with surface changes in language that would support a learner-centered hypothesis: There is no concordant change in word length across corpora nor is there a preference for producing words with earlier age of acquisition. In addition, we also find no evidence that this change is a function of semantic bleaching: In a comparison of two different concreteness norms taken 45 years apart, we find no systematic change in word concreteness. Finally, we test the crowding hypothesis directly by comparing approximately 1.5 million tweets across the 50 US states and find a correlation between population density and tweet concreteness, which is not explained by state IQ. The results demonstrate both how Big Data can be used to discriminate among alternative hypothesis as well as how language may be changing over time in response to the rising tide of information.

Language has changed. Consider the following example from *The Scarlet Letter*, written by Nathaniel Hawthorne and originally published in 1850 (Hawthorne, 2004):

> Children have always a sympathy in the agitations of those connected with them; always, especially, a sense of any trouble or impending revolution, of whatever kind, in domestic circumstances; and therefore Pearl, who was the gem on her mother's unquiet bosom, betrayed, by the very dance of her spirits, the emotions which none could detect in the marble passiveness of Hester's brow.

Compare this with a more recent example of writing from David Sidaris's *Barrel Fever*, published in 1994 (Sidaris, 1994):

> If you're looking for sympathy you'll find it between shit and syphilis in the dictionary.

Both of these examples contain words that readers are likely to recognize. Both allude to fairly abstract concepts in relation to the word *sympathy*. However, they use dramatically different words to do so. Yet, with but a few examples, there is always the risk that one is cherry picking. Compare, as another example, the following two quotations, both published in *Nature*, but more than 100 years apart:

> When so little is really known about evolution, even in the sphere of organic matter, where this grand principle was first prominently brought before our notice, it may perhaps seem premature to pursue its action further back in the history of the universe. (Blanshard, 1873)

> Each sex is part of the environment of the other sex. This may lead to perpetual coevolution between the sexes, when adaptation by one sex reduces fitness of the other. (Rice, 1996)

The difference in language between these examples hints at something systematic, with the more recent examples appearing more efficient—they are perhaps more conceptually clear and perhaps more memorable, arguably without any loss of content or fuzzying of the intended meaning. One could consider this a kind of conceptual evolution in the language. Still, these are but a few examples and the change is difficult to quantify. To characterize conceptual change in language over the past several hundred years would require reading thousands or millions of books and attempting to precipitate out change in a psycholinguistic variable capable of capturing this underlying change in conceptual efficiency. The computational and data resources associated with "Big Data" make this work possible. In recent years millions of words of natural language have been digitized and made available to researchers, and large scale psycholinguistic norms have been published using online crowdsourcing to rate thousands of words. In this chapter, we combine a wide collection of such resources—books, newspapers, magazines, public speeches, tweets, and psycholinguistic norms—to investigate the recent psycholinguistic evolution of American English.

Language Change

It is well established that languages change over time (Labov, 1980; Lieberman, Michel, Jackson, Tang, & Nowak, 2007; Michel, et al., 2011), and this evolution has been characterized at both long time scales, on the order of thousands of years (Pagel, et al., 2007; Lupyan, & Dale, 2010), as well as shorter time scales, on the order of generations (Labov, 1980; Michel, et al., 2011). Word forms change, such as the transition from Old English *lufie* to present day *love*. Words are replaced by other words, such as replacement of Old English *wulcen* by the Scandinavian word *sky* (Romaine, et al., 1992). Words change their grammatical value, such as the word *going*, which once strictly referred to physical movement from one place to another and now refers to the future, as in "They are going to change their policy". And words often change in more than one way at the same time, as in the Proto-Indo-European word *kap*, which may have originally meant "to seize" (as in the Latin root *cap*, from "capture" and "captive"), but now hides in plain sight in English as the word *have* (Deutscher, 2010).

The factors influencing which words change and why has recently seen an upsurge of interest outside traditional linguistics based on the availability of large databases of language. Using this approach, researchers have isolated a variety of factors that influence language change. For example, the average half-life of a word is between 2,000 and 4,000 years (i.e. the time at which half of all words are replaced by a new non-cognate word), but the half-life of a word is prolonged for high-frequency words (Pagel, et al., 2007) and for words with earlier age of acquisition (Monaghan, 2014). Languages are also susceptible to losing grammatically obligatory word affixes, called morphological complexity. This varies across languages in predictable ways as a function of the number of speakers. English, compared to most other languages, has very little morphological complexity. For example, English speakers add the suffix *-ed* to regular verbs to signal the past tense. Navajo, on the other hand, adds obligatory morphemes to indicate a wide variety of information about past action, for example, whether it occurred among groups, was repeated, or changed over time (Young, Morgan, & Midgette, 1992). In a comparison of more than 2,000 languages, languages with more speakers were found to have less morphological complexity than languages with fewer speakers (Lupyan, & Dale, 2010). All of the above examples indicate that potential selective pressures such as frequency of shared language use can preserve language features and that these can be characterized, for example, in much the same way as genetic evolution (compare, e.g. Pagel, et al., 2007; Duret & Mouchiroud, 2000).

The hypothesis that languages evolve in response to selective pressures provided by the social environment in which they are spoken is called the *linguistic niche hypothesis* (Lupyan, & Dale, 2010). For example, Lupyan, & Dale (2010) interpret the above changes in morphological complexity as evidence consistent with a selective force introduced by an influx of adult language learners among languages with more speakers. This shares much in common with other learner-centered accounts of language change (Trudgill, 2002; McWhorter, 2007). Our goal in the present work is to extend the linguistic niche hypothesis by developing theory around another potential source of social selection, and in particular one that is likely to dominantly influence patterns of language use.

Conceptual Crowding

Cognitive performance is intimately connected with our capacity to process, understand, and recall new information. One of the principal features of cognitive information is that it can experience crowding. An example is provided by list-length effects, where items on a longer list are remembered less well than items on shorter lists (Ratcliff, Clark, & Shiffrin, 1990). Similar information crowding may also be taking place at the level of our day to day experience where the production and exposure to information has seen unprecedented growth (Varian & Lyman, 2000; Eppler & Mengis, 2004). Such environments should create highly competitive environments for information, especially among languages with many speakers, who each represent potential competitors in the language marketplace. This competition in information creates information markets, which have recently been associated with *attention economies* (Hansen & Haas, 2001; Davenport & Beck, 2001). That is, attention represents a limiting resource in information rich environments. Attention economies can drive evolutionary change in communication in the same way that noise drives change in other communication systems. For example, acoustic crowding associated with bird song is known to influence communication strategies in bird species, in ways that make songs easier to discriminate (Luther, 2009; Grant & Grant, 2010). Similar to a global cocktail party problem—where communicators compete for the attention of listeners—conceptual information may also compete for attention, both at the source—when it is spoken—as well as later, when it is being retrieved.

To formalize the above argument, consider that crowding mimics the inclusion of noise in information theoretic accounts of signal transduction (Shannon, & Weaver, 1949). This is often discussed in terms of the signal to noise ratio, $\frac{S}{N}$. From the perspective of an information producer, other producer's messages are noise, and as these increase our own messages are transmitted less efficiently.

How should cognitive systems adapt to this situation? Simon provided an answer for receivers: "a wealth of information creates a poverty of attention and a need to allocate that attention efficiently" (Simon, 1971: 40–41). One way to accomplish this is to order information sources in proportion to their value to the receiver. Here, people should learn to tune out (i.e. inhibit) unwanted messages and pay more attention to those that represent real conceptual value. This is the basis of algorithmic approaches to filtering junk e-mail (e.g. Sahami, Dumais, Heckerman,

& Horvitz, 1998). Thus, conceptually ambiguous messages may be tuned out in the crowded information market. However, message senders can also respond to crowded information markets by changing the message. As language adapts to the needs of its users (Christiansen & Chater, 2008), language should also change in response to information crowding. In particular, language should evolve to increase signal strength, S.

How can languages increase S? Shannon's mathematical account of information theory is not so helpful here. This is because S must necessarily be adapted to the characteristics of the receiver. In the case of language, this means that messages which are most likely to be processed effectively are messages that are more rapidly understood and later recalled. Thus, one way to increase S is to reduce the conceptual length of a message.

An Illustrative Example: Optimal Conceptual Length

To provide a quantitative illustration of the impact of noise on the conceptual evolution of language, consider the following example. Let a message of bit size, b, indicate truth information about the world, which provides information about the locations of resource targets. Here we take b to represent the conceptual length of a message. Suppose that the number of targets, n_b, identified in a message of a given length is $n_b = b^\lambda$. When $\lambda > 1$ longer messages contain more information per conceptual bit than short messages. If the error rate per bit is f, then the probability that a message is received without error is $\omega = (1 - f)^b$.

Assume that transmitters repeat messages about targets until they are successfully received and then transmit locations of additional targets. With a constant bit rate, the target acquisition rate for messages of length b is

$$R = \frac{b^\lambda \omega}{b} \tag{1}$$

which is the product of the message value in targets and the rate of successful signal transmission. $R(b)$ is maximized when

$$b^* = -\frac{\lambda - 1}{ln(1 - f)} \tag{2}$$

Formally: $b^* \to 0$ *as* $f \to 1$.

Thus, as noise increases signal lengths should become smaller. Figure 12.1 shows this relationship for a variety of error rates. This is a general finding that is not constrained to words, but to messages and the time it takes to process them. It is similar to the inverse relationship between channel capacity and noise in Shannon's noisy-channel encoding theorem (MacKay, 2003), except here we are concerned with optimal message length when message interruption leads to complete message failure.

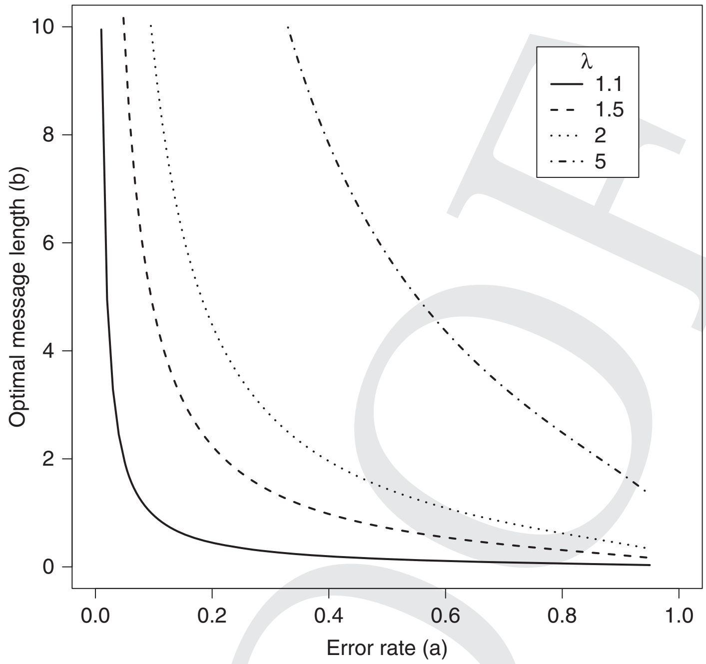

FIGURE 12.1 Optimal message length as a function of error rate.

Importantly, the result points to how conceptual information in communication should respond to noise in the form of information crowding. As a selective force, crowding should enhance the conceptual efficiency of language by reducing the conceptual length of messages. In the absence of crowding, this selective force is reduced and languages may respond to other influences, such as signalling information about the speaker or providing details that offer more information with respect to the intended message. In crowded environments, the cost of this additional information may jeapordize the message in its entirety.

Surface Versus Conceptual Complexity

The above example leaves ambiguous what is precisely meant by conceptual length. That is, the processing required to render a message into its cognitive output depends on multiple aspects. A receiver of a message must not only veridically identify the particular surface forms that have been transmitted, but also interpret and understand the conceptual content of the message. Thus, the amount of processing required depends on message form and content. A reader will have less work to do to determine who has done what to which object when the writer writes "Dr Estes dropped his morning banana on the desk in Dr Wade's office" than when he writes "The professor's mishandling of his fruit influenced his colleague's work". This is true even for a reader with enough knowledge to disambiguate the references, because recomposing a message from its parts requires cognitive processing and more so when those parts are ambiguous (Murphy & Smith, 1982; Rosch, Mervis, Gray, Johnson, & Boyesbraem, 1976).

In relation to the above formal example, messages that require more cognitive processing are more vulnerable to incomplete transmission. First, each unit of processing work is vulnerable to intrinsic failure; the greater the amount of work, the higher the probability of failure. All else being equal, this will result in more complex messages being lost. Second, as discussed above, there is competition for the receiver's resources. New messages may arrive during the processing of older messages, and interrupt their transmission.

Messages may also be more complex than their payloads require. Take for example the following sentence, which won the *Bad Writing Contest* in 1997:

> The move from a structuralist account in which capital is understood to structure social relations in relatively homologous ways to a view of hegemony in which power relations are subject to repetition, convergence, and rearticulation brought the question of temporality into the thinking of structure, and marked a shift from a form of Althusserian theory that takes structural totalities as theoretical objects to one in which the insights into the contingent possibility of structure inaugurate a renewed conception of hegemony as bound up with the contingent sites and strategies of the rearticulation of power. (Butler, 1997)

Regardless of how profound the sentence may be, many have found it to be most valuable as a lesson in maximizing words per concept (e.g. Pinker, 2014).

In the Dr. Estes example above, using less specific terms such as "fruit" when "banana" is meant, increases the conceptual complexity of the message without changing the payload, and without a gain in surface complexity. Writers can eliminate such inefficiencies, but they may not do so unless pressured because the cost to language users of being more specific is non-zero. G. K. Zipf referred to a similar balance between reductive and expansive forces in language:

> ... whenever a person uses words to convey meanings he will automatically try to get his ideas across most efficiently by seeking a balance between the economy of a small wieldy vocabulary of more general reference on the one hand, and the economy of a larger one of more precise reference on the other, with the result that the vocabulary of *n* different words in his resulting flow of speech will represent a *vocabulary balance* between our theoretical forces of unification and diversification. (Zipf, 1949: 22)

Thus, crowding is but one force that language users must accomodate for.

The formal bases for the costs of surface and conceptual complexity are analogous, but they are subject to different trade-offs. At its simplest, surface complexity corresponds to physical length (e.g. word length), suggesting a cost for longer words. However, longer words are also more phonologically isolated, such that they are effectively error-correcting codes: If a long word is disrupted slightly, it is still unlikely to be confused with another word, but if a short word suffers distortion, the outcome can be more easily confused with another word. Conceptual length is not necessarily associated with physical length, because it relies on higher-order cognitive processing capacities involved in message retrieval and production. In the next section we expand further on this by relating conceptual length to historical change in language.

Conceptual Efficiency and Concreteness

The properties of language that are best associated with conceptual efficiency—being more rapidly understood and later recalled—have been extensively studied by psycholinguists. In particular, concreteness is well marked by its ability to enhance these processing capacities. Concreteness refers to a word's ability to make specific and definite reference to particular objects. Among psycholinguistic variables, the range and depth of concreteness on cognitive processing is easily among the most profound. Paivio's dual-coding theory (Paivio, 1971), which proposed both a visual and verbal contribution to linguistic information, led to years of research showing that concrete words had a memory advantage in recall tasks (Paivio, Walsh, & Bons, 1994; Romani, McAlpine, & Martin, 2008; Fliessbach, Weis, Klaver, Elger, & Weber, 2006; Miller & Roodenrys, 2009). This initial research has since led to numerous studies articulating the influence of concreteness as an important psycholinguistic variable. A Google Scholar search of "concreteness" and "linguistics" finds approximately 30,000 articles that contain both terms, with approximately 2,000 published in the last year.

The breadth of this research is compelling. Concrete language is perceived as more truthful (Hansen & Wanke, 2010), and it is more interesting and easier to comprehend (Sadoski, 2001). Concrete words are recognized more quickly in lexical decision tasks than more abstract words (James, 1975). Concrete words show an advantage in bilingual translation and novel word learning (Tokowicz & Kroll, 2007; Kaushanskaya & Rechtzigel, 2012), and concrete words are more readily learned by both second and first language learners (De Groot & Keijzer, 2000). In addition, concrete and abstract words are processed differently in the brain (Adorni & Proverbio, 2012; Huang, Lee, & Federmeier, 2010).

Multiple explanations for the importance of concreteness have been proposed, ranging from the imagibility of words (Paivio, 1971), to their contextual availability (Schwanenflugel, Harnishfeger, & Stowe, 1988), to a more recent account based on emotional valance (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011).

The important point here is that each of these theories acknowledges the powerful role that concreteness plays in facilitating linguistic processing.

The wealth of evidence on the value of concreteness in language presents a problem. Why should words ever become abstract? The assumption in the mathematical example above provides what we feel is the most likely explanation—abstract words, by the nature of their generality, provide information about broader classes of phenomena. The word *fruit*, being more abstract than items subordinate to that category, e.g. *apple*, can efficiently communicate information about categories of truth that would otherwise take many individual instances of more concrete examples—"fruit can prevent scurvy when eaten on long ocean voyages". Or to take another example, the word *essentially* is one of the most abstract words in concreteness norms (see Brysbaert, et al., 2013), however the sentence "My office hours are essentially on Wednesday at noon" provides a degree of probabilistic hedging that "My office hours are on Wednesday at noon" does not.

Besides the theory of crowding proposed above, we know of no prior theories that speak directly to evolutionary changes in concreteness at the level of word distributions. Nonetheless, some evidence of cultural change may speak indirectly to this issue. The most prominent is associated with an explanation for the Flynn effect. The Flynn effect is the observation that intelligence scores, associated with both crystallized and fluid intelligence, have risen steadily from approximately the middle of the last decade (Flynn, 2012). Flynn noted that nutrition has failed to explain the observed effects, and that in the absence of evidence for other biological theories, more cognitive theories have risen to the foreground (Flynn, 2012). In particular, numerous theories speak to an increase in computational, symbolic, and potentially more abstract processing abilities (Greenfield, 1998; Fox, & Mitchum, 2013). One implication of knowledge-based accounts is that language may change its composition to reflect our capacity for more abstract processing, and thus show an increase in abstract words.

However, the causal arrow may point in the other direction. An increase in concrete language may enhance our individual capacity to process complex information. By this interpretation, language should have become more concrete, and where it is concrete people should tend to learn and process more about it (Sadoski, 2001).

The Rise in Concreteness

We combined multi-billion word diachronic language corpora (e.g. the Google Ngram corpora and the Corpus of Historical American English) with a recent collection of conctreness norms composed of concreteness ratings for approximately 40,000 words (Brysbaert, et al., 2013). The Google Ngram corpus of American English provides a collection of over 355 billion words published in books digitized by Google (Michel, et al., 2011). We limit our analysis to years following 1800 to ensure appropriate representation. The Corpus of Historical American English, collected independently of the Google Ngram corpus, represents a balanced and representative corpus of American English containing more than 400 million words of text from 1810 to 1990, by decade, and composed of newspaper and magazine articles (Davies, 2009). In addition, we included presidential inaugural addresses (Bird, Klein, & Loper, 2009), with measures computed from the entire text of each speech. All of the data sources are freely available online.

We tracked changes in concreteness over time by computing a measure of average concreteness. The concreteness index, C_y, for each year, y, in a corpus was computed as follows,

$$C_y = \sum_i^n c_i p_{i,y}$$

where c_i is the concreteness for word i as found in the Brysbaert et al. concreteness norms and p_i is the proportion of word i in year y. The proportion is computed only over the n words in the concreteness norms, or the appropriate comparison set (as described in the figure caption). We also computed concreteness on a per document basis, as opposed to per word, with similar results.

As shown in Figure 12.2, we found a steady rise in concreteness across multiple corpora, including books (Google ngrams), newspapers and magazines (the Corpus of Historical American English), and presidential speeches. The Google Ngrams also provide a corpus based on English fiction, which shows the same pattern, with a slightly more dramatic rise in concreteness from approximately 2.35 in 1800 to 2.57 in 2000 (data not shown).

We also found that changes in concreteness occurred within word classes and were therefore not strictly due to changes in syntax (e.g. by a reduction in the use of articles). Figure 12.3 shows that, over the same time period, concreteness increases within nouns, verbs, and prepositions. Together these findings show that the change is not only systematic but systemic, permeating many different aspects of the language.

This observation is consistent with a number of hypotheses, including crowding and the influence of second language learners. In what follows, we examine a number of these hypotheses in an effort to discriminate among what is, at least initially, a wide set of possibilities.

Semantic Bleaching

It has been proposed that word evolution follows a fairly predictable pattern over time, from specific to abstract. This takes a variety of forms including semantic bleaching, desemanticization, and grammaticalization (Hopper & Traugott, 2003; Aitchison & Lewis, 2003). An example is the word *very*, which derives from the French word *vrai*. In French, *vrai* did and still does mean "true". In Middle English,

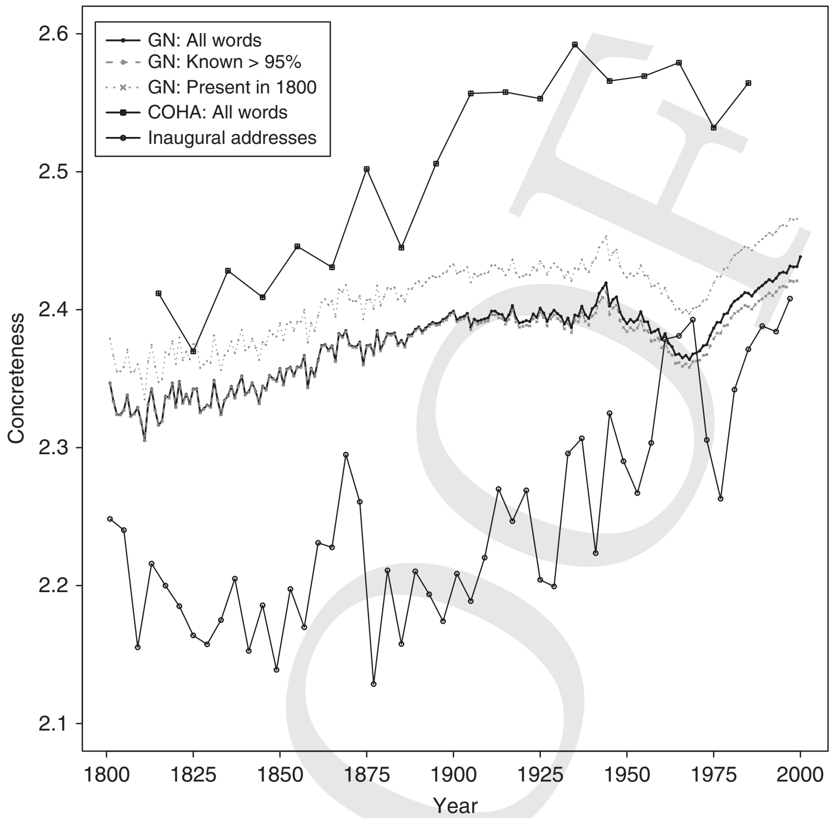

FIGURE 12.2 Concreteness in the Google Ngrams (GN), Corpus of Historical American English (COHA), and presidential inaugural addresses. Also shown is the Google Ngram corpus using only words that were present in the 1800 lexicon or only words that were known by more than 95 percent of participants in the concreteness norms (Brysbaert, et al., 2013). Pearson correlations with time (95 percent confidence interval): Google Ngrams, $r = 0.86$ (0.82, 0.89); COHA, $r = 0.89$ (0.73, 0.96); Inaugural addresses, $r = 0.77$ (0.63, 0.86). Figure taken from Hills & Adelman, 2015.

the word *very* also meant “true”, as in a *very knight*, meaning a real knight. However, over time the word became a means to emphasize the strength of a relationship, in a probabilistic way. For example, nowadays we say that something is *very true*, meaning that there are degrees of truth and this particular thing may have more of it than others.

Although semantic bleaching is proposed to be unidirectional, this is not without debate (Hollmann, 2009). Moreover, it is certainly the case that not all diachronic linguistic patterns are associated with loss of concreteness. Metonymy

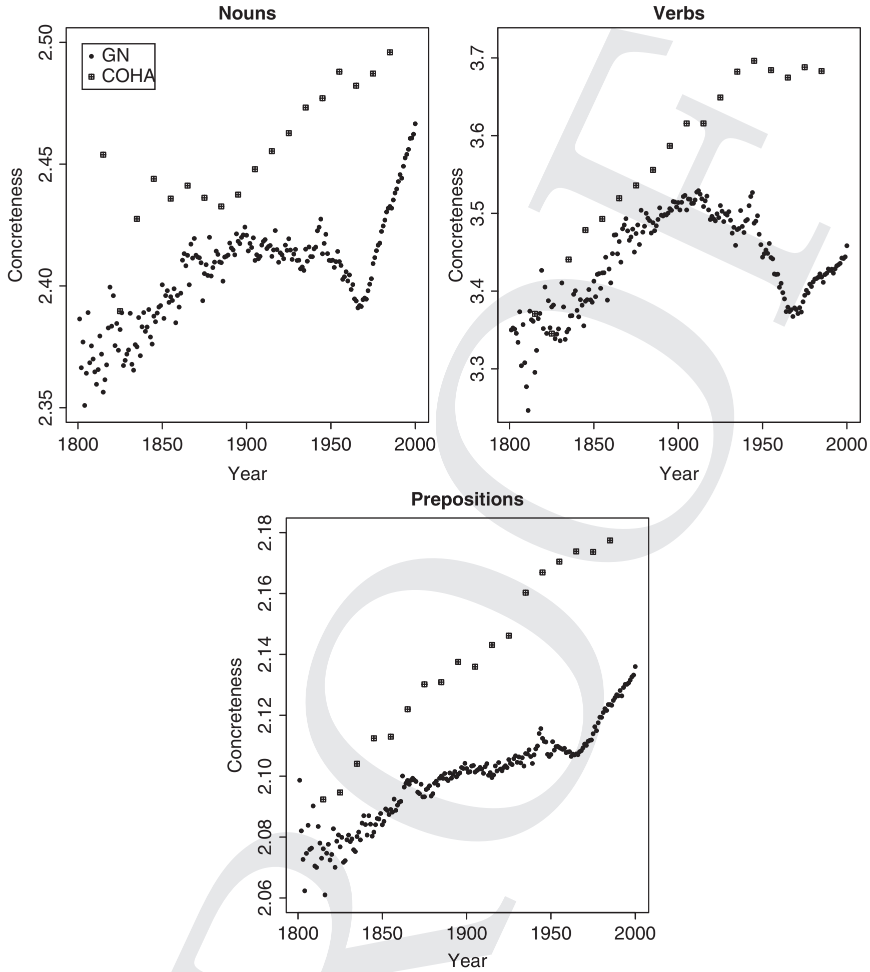

FIGURE 12.3 Changes in concreteness within word classes. Figure taken from Hills & Adelman, 2015.

is a common figure of speech where some specific aspect of an abstract concept is used in its place, as in the phrases *the pen is mightier than the sword* and *the Pentagon*.

However, if bleaching were sufficiently strong, it could explain the observed rise in concreteness. Language could have been perceived as more concrete by its users at the time, but appear less concrete now because the norms are based on present day perceptions of the word concreteness.

Concreteness norms were not collected as far back as the 1800s. However, the existence of the Paivio norms (Paivio, et al., 1968) provides a 45-year window of

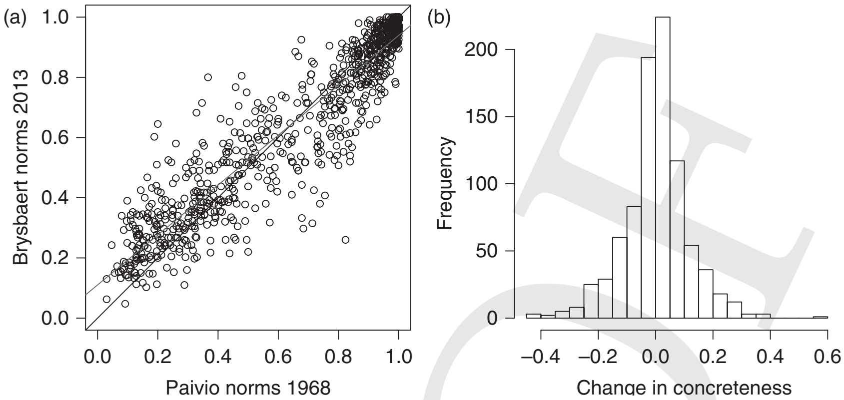

FIGURE 12.4 Comparison of the Paivio and Brysbeart norms. The Paivio concreteness norms (Paivio, et al., 1968) consist of 925 nouns, collected in the laboratory and using a 7-point scale. The Brysbaert norms were collected on a 5-point scale. Both norms are normalized to be between zero and one. (a) Shows the change in concreteness over the 45-year span between collection. (b) Shows the histogram of concreteness differences per word. Figure taken from Hills & Adelman, 2015.

comparison with the Brysbaert norms and provides the only basis we know of for a quantitative test of semantic bleaching. Normalizing the ratings for both shows that there are no systematic changes in word concreteness over the approximately 900 words used for comparison (Figure 12.4). The median change is centered around zero and a paired t-test finds no significant difference in concreteness ($t(876) = -0.79$, $p = 0.45$). This suggests that a systematic loss of concreteness is unlikely to explain the apparent rise in concreteness we see in the data.

These results also provide a large scale approach to better understanding the unidirectionality of semantic bleaching, which to our knowledge has not been possible in the past. As a preliminary step in that direction, in Table 12.1 we provide the ten words that have increased or decreased the most in concreteness over the last 45 years. Importantly, the words from each side of the distribution offer an effective demonstration that semantic bleaching may be mirrored by an equally powerful semantic enrichment. A *dreamer* may have become more concrete—but the *devil*, although he may have been in the details in the past, resides more in abstraction today.

Reductions in Surface Complexity

The conceptual evolution described above is consistent with our information crowding hypothesis, but it is also consistent with a learner-centered hypothesis (e.g. Lupyan, & Dale, 2010). Indeed, concreteness may represent one of many changes reflecting an overall adaptation in English to selective forces driven

TABLE 12.1 Lists of words that have increased or decreased the most in concreteness between the Paivio and Brysbaert concreteness norms. Positive values indicate an increase in concreteness over time.

Word	*Change*	*Word*	*Change*
sovereign	0.56	facility	−0.45
plain	0.40	month	−0.43
dreamer	0.39	devil	−0.41
originator	0.36	competition	−0.37
master	0.35	death	−0.35
outsider	0.35	vacuum	−0.33
habitation	0.32	demon	−0.32
antitoxin	0.30	panic	−0.31
connoisseur	0.30	length	−0.31
captive	0.29	sensation	−0.30

by language learners. A hypothesis based on language learners—assuming it is operating over the timescale in which we investigate American English—should also predict surface simplification in the language. Surface simplification would include features such as shorter word length and preference for words with earlier age of acquisition. The absence of these features changing alongside concreteness in no way discounts prior evidence for learner-centered change in language more generally. However, it would indicate that the changes driving concreteness in American English may not be learner-centered or a result of language speakers becoming more childlike, but may instead be driven by factors more specifically associated with conceptual clarity induced by crowding.

Word Length

More concrete words tend to be phonologically and orthographically shorter. Among the words in the Brysbaert norms (Brysbaert, et al., 2013), the correlation between concreteness and word length is β =-0.40, $p < 0.001$. If the selective forces driving concreteness are more directly driven by preferences for shorter words, then word length should change in tandem with concreteness. However, Figure 12.5 shows that the general trends found in concreteness are not preserved across corpora in relation to word length. In general, word length does not change much across the three corpora until the 1900s, and then the direction of change appears to depend on the corpora. Words in presidential speeches get shorter, while words in books tend to get longer. Words in newspapers and magazines, on the

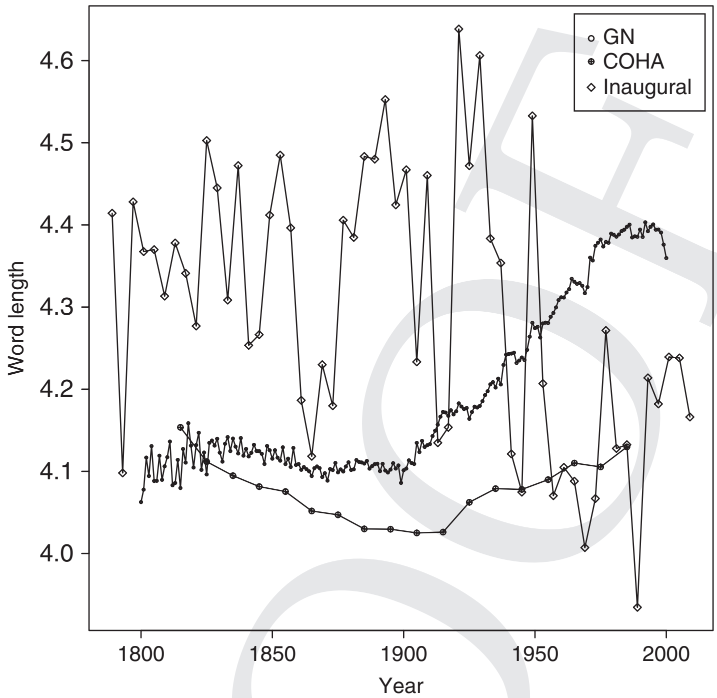

FIGURE 12.5 Changes in average word length over time for words in the Google Ngrams (GN), Corpus of Historical American English (COHA), and presidential inaugural addresses.

other hand, first show a trend towards reduced length but then increase in length, but only up to approximately the point they were in 1800.

One can also look at concreteness with words of a given length, and ask if the rise in concreteness is independent of word length. Figure 12.6 shows that this is largely the case. Although words of one, two, or three characters in length show roughly no change in concreteness over time, words of four or more characters consistently show a rise in concreteness over time ($ps < 0.001$).

Additional evidence of the independence between concreteness and word length is found in Figure 12.7, which shows that within a range of concreteness words tend to grow in length, especially following the 1900s. This is also mirrored by an increase in word length across the full corpus. This would appear to be counter to a potential selective force imposed by language learners. In sum, changes in concreteness do not appear to be driven by changes in word length—on the contrary, concreteness appears to rise despite an overall trend towards longer words.

"9781138791923C12" — 2016/8/22 — 19:44 — page 285 — #16

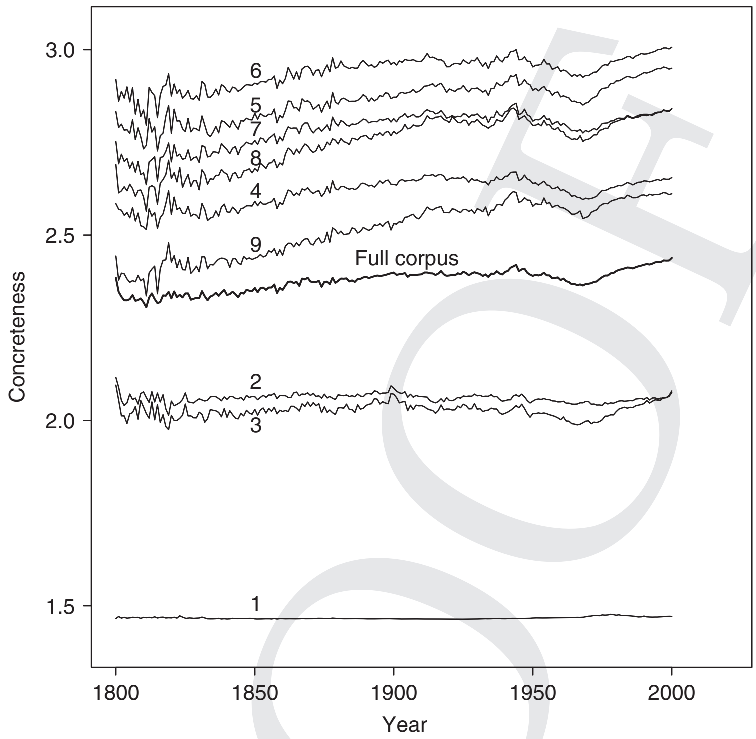

FIGURE 12.6 Changes in concreteness within words of a given length in characters. Data are taken from the Google Ngrams.

Age of Acquisition

Age of acquisition provides an additional, and possibly more direct, measure of evidence for a learner-centered hypothesis. In a comparision of the 23,858 words that are shared between the Brysbaert concreteness norms and the Kuperman age of acquisition norms (Kuperman, et al., 2012), age of acquisition is correlated with concreteness, $\beta = -0.35, p < 0.001$. Moreover, previous work has established that words with earlier age of acquisition are more resilient to age-related decline and are retrieved more quickly in lexical decision tasks than words acquired later in life (Juhasz, 2005; Hodgson & Ellis, 1998). If language change in American English is driven by the influence of language learners—who may only show partial learning—or the influence of an aging population—who produce earlier acquired words preferentially—then the language as a whole may move towards words of earlier age of acquisition.

To evaluate changes in preference for words of earlier acquisition over time, we used the Kuperman age of acquisition norms (Kuperman, et al., 2012) to compute

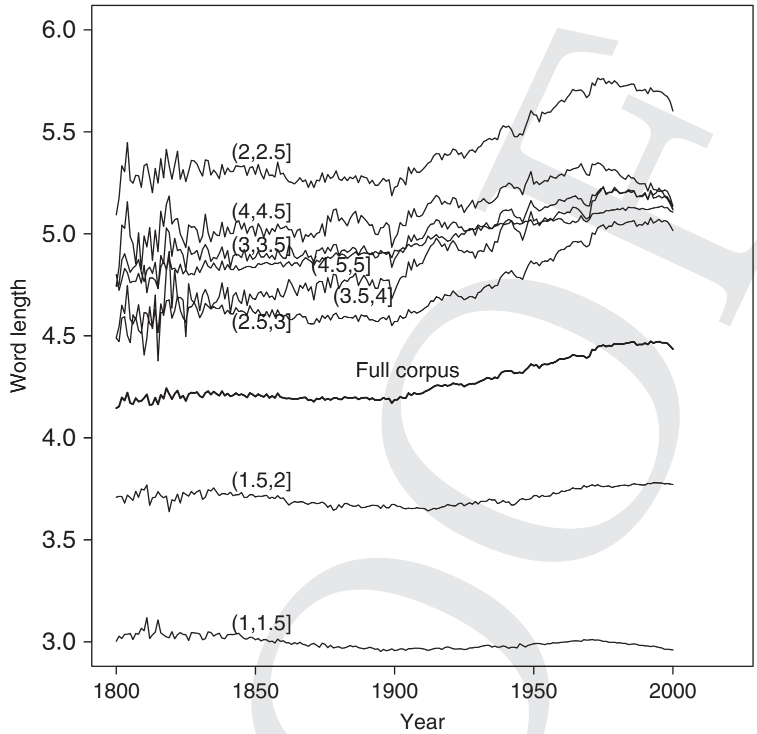

FIGURE 12.7 Changes in word length within narrow ranges of concreteness. Data are taken from the Google Ngrams.

a weighted age of acquisition value for each corpora as was done for concreteness. Figure 12.8 shows that age of acquisition tends to follow a similar pattern as that found for word length, but not concreteness. Changes in age of acquisition and word length are also highly correlated across the three corpora (Google Ngram: $\beta = 0.96$, $p < 0.001$; COHA: $\beta = 0.66$, $p < 0.001$; inaugural addresses: $\beta = 0.95$, $p < 0.001$). On the other hand age of acquisition is not well correlated with changes in concreteness (e.g. Google Ngram: $\beta = 0.33$, $p < 0.001$).

Discussion of the Absence of Reduction in Surface Complexity

The above evidence for an absence of systematic changes in word length and age of acquisition suggest that the observed changes in concreteness are not the result of factors that might also lead to surface simplication. Specifically, the evidence suggests that concreteness is not being driven by second language

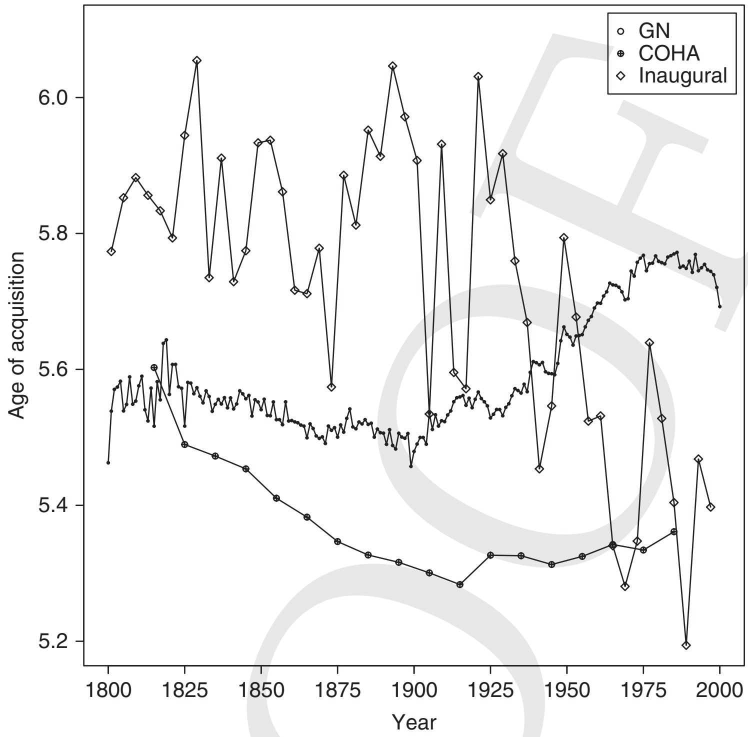

FIGURE 12.8 Changes in average age of acquisition for words in the Google Ngrams (GN), Corpus of Historical American English (COHA), and presidential inaugural addresses.

learners, who we would predict would also show a preference for shorter words and words with earlier age of acquisition. Furthermore, the results also suggest that the change in concreteness is not being driven by a rising tide of more children or lower IQ individuals entering the language market. Again, if this were the case, we would expect language to also show systematic changes in surface complexity. In the next section we examine more directly the relationship between crowding and concrete language by looking at another source of Big Data: Twitter.

Population Density and Concreteness in US States

As a final look into the relationship between crowding and concreteness, we investigated the concreteness of tweets across the 50 US states. Here our prediction is that states with a higher population density should also produce more concrete language. To investigate this, from 24 February 2014 till 8 April 2014, we

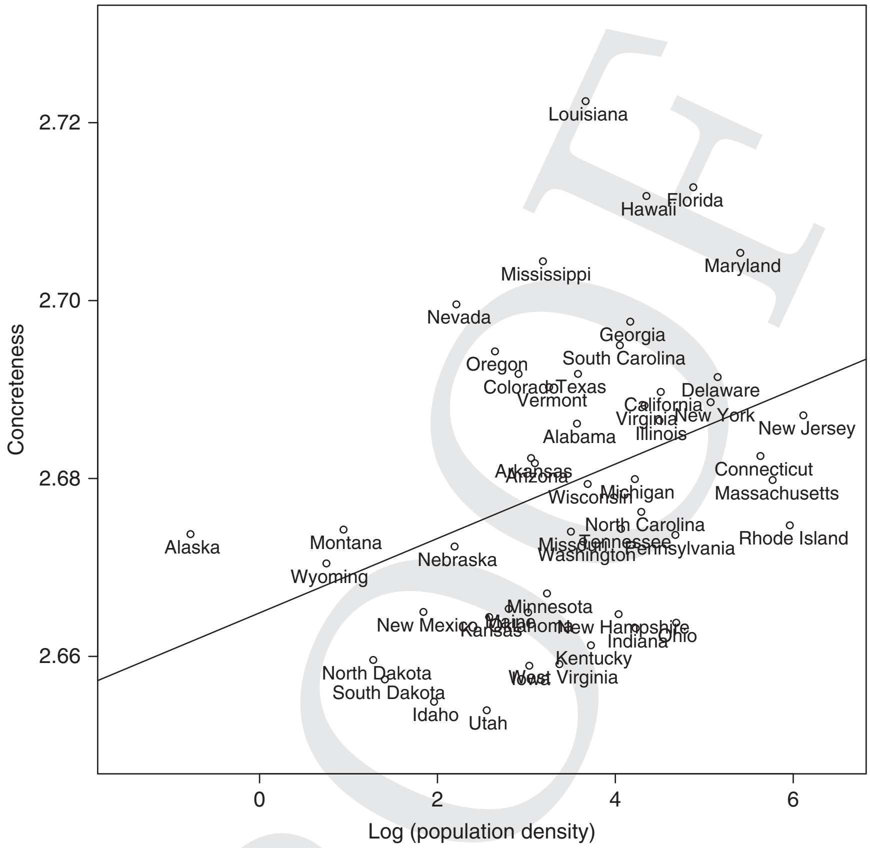

FIGURE 12.9 Concreteness in language increases with the population density across US states. The data are taken from approximately 1.5 million tweets, with 30,000 tweets per state.

collected 66,348,615 tweets, made within 50 states of the USA, using Twitter's streaming API. The collected tweets exclude retweets (i.e. repetition of tweets previously made). The number of collected tweets varies between the states from 39,397 (Wyoming) to 8,009,114 (California), and thus to achieve similarity in measurement accuracy, we randomly sampled 30,000 tweets from each state. Then after removing hash tags, non-ascii characters, and punctuation marks (e.g."), we calculated the concreteness for each tweet and then mean averaged these for each state.

Figure 12.9 shows the relationship between log population density and tweet concreteness for states ($\beta = 0.36,\ p < 0.01$). There is a clear pattern of rising concreteness with population density. There are many potential confounds here, as styles of writing (e.g. syntax and tweet length) may change across states. However, as we note above, concreteness is but one of many ways that conceptual efficiency may change and thus we see it as an indicator, which may in turn be driven by other aspects of language use. One factor that is unlikely to be influenced by crowding, however, is IQ, which may in turn be associated with concreteness, as we note in the introduction. In our data, IQ is inversely correlated with concreteness ($\beta = -0.003$, $p < 0.001$), but this may not be particulary surprising as the McDaniel (2006) measure is based partly on reading comprehension. However, the relationship between concreteness and population density is preserved after partialing out the variance accounted for by changes in IQ (McDaniel, 2006), with population density still making up approximately 12 percent of the variance ($p < 0.01$).

Conclusions

Culture is a marketplace for competing ideas. This leads to the prediction that any broad medium of communication should evolve over time to have certain properties that facilitate communication. Moreover, certain aspects of these signals should be enhanced as competition (i.e. crowding) increases. In particular, aspects of information that increase speed of processing and memorability should be favored as information markets become more crowded. As we have shown, concreteness facilitates these cognitive demands and has risen systematically in American English for at least the last 200 years. We have also shown that these changes are not consistent with a learner-centered hypothesis, because we would then expect additional changes in language associated with a reduction in surface complexity, such as reduced word length and preference for words with earlier age of acquisition, which we do not observe. The lack of evidence for these changes also indicates that the change in concreteness is not due to a general simplifying of the language, which one might predict if language were being influenced by, for example, a younger age of entry into the language marketplace or a general dumbing down of the language.

The work we present here is preliminary in many respects. We have taken a bird's eye view of language and focused on psycholinguistic change, but these necessarily require some assumptions on our part and do not focus on other factors in language change. It is very likely that there are other changes in writing and speech conventions that one could document. To see how these align with our present investigation, one would also need to investigate the causes of these changes. If writing was once meant to provide additional information about the intelligence of the author, this may have been lost in modern language—but the natural question is why? When there are but few authors, the authors may compete along different dimensions than when there are many, and conventions may change accordingly.

The present work demonstrates the capacity for data analytic approaches to language change that can discriminate among alternative hypotheses and even combine data from multiple sources to better inform hypotheses. Naturally, we also hope this work leads to future questions and research on the influence of concreteness and language evolution. In particular, we find it particularly intriguing to ask if the rise in concrete language may be associated with the rise in IQ associated with the Flynn effect (Flynn, 2012)? Compared with the writing of several hundred years ago, the examples we provide in the introduction suggest that today's writing is more succinct, often to the point of being terse. It is difficult to deny the comparative ease with which modern language conveys its message. Indeed, we suspect that more memorable language (such as aphorisms) share a similar property of making their point clearly and efficiently.

Additionally, the research we present here also poses questions about the influence of competition in language. Is language produced in a competitive environment more memorable in general, or is the increased memorability for some passages just a consequence of a larger degree of variance among the language produced? If the former, this may suggest that something about competitive language environments facilitates the production of more memorable messages, and that this is something that humans are potentially aware of and capable of modulating. Such a capacity would explain the enhanced memorability of Facebook status updates relative to other forms of language (Mickes, Darby, Hwe, Bajic, Warker, Harris, & Christenfeld, 2013). If such competition is driving language, and language change maintains its current course, we may all be speaking Twitterese in the next 100 to 500 years (compare the y-axis on Figure 12.2 and Figure 12.9). Finally, this work may also provide applications in relation to producing more memorable information in learning environments, for example, via a mechanism for concretizing text or competitive writing. Although we recognize that these questions are speculative, we hope that this work provides some inspiration for their further investigation.

References

Adorni, R., & Proverbio, A. M. (2012). The neural manifestation of the word concreteness effect: an electrical neuroimaging study. *Neuropsychologia, 50,* 880–891.

Aitchison, J., & Lewis, D. M. (2003). Polysemy and bleaching. In B. Nerlich, Z. Todd, V. Hermann, & D. D. Clarke (Eds.), *Polysemy: flexible patterns of meaning in mind and language* (pp. 253–265). Berlin: Walter de Gruyter.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python*. Sebastopol, CA, USA:O'Reilly.

Blanshard, C. (1873). Evolution as applied to the chemical elements. *Nature, 9,* 6–8.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2013). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46,* 1–8.

Butler, J. (1997). Further reflections on conversations of our time. *Diacritics, 27,* 13–15.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31,* 489–509.

Davenport, T. H., & Beck, J. C. (2001). *The attention economy: Understanding the new currency of business.* Brighton, MA, USA Harvard Business Press.

Davies, M. (2009). The 385+ million word corpus of contemporary American English (1990– 2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics, 14,* 159–190.

De Groot, A., & Keijzer, R. (2000). What is hard to learn is easy to forget: the roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning, 50,* 1–56.

Deutscher, G. (2010). *The unfolding of language.* London: Random House.

Duret, L., & Mouchiroud, D. (2000). Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Molecular Biology and Evolution, 17,* 68–670.

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: a review of literature from organization science, accounting, marketing, mis, and related disciplines. *The Information Society, 20,* 325–344.

Fliessbach, K., Weis, S., Klaver, P., Elger, C., & Weber, B. (2006). The effect of word concreteness on recognition memory. *Neuroimage, 32,* 1413–1421.

Flynn, J. R. (2012). *Are we getting smarter?: Rising IQ in the twenty-first century*. Cambridge, UK: Cambridge University Press.

Fox, M. C., & Mitchum, A. L. (2013). A knowledge-based theory of rising scores on 'culture- free' tests. *Journal of Experimental Psychology: General, 142,* 979–1000.

Grant, B. R., & Grant, P. R. (2010). Songs of Darwin's finches diverge when a new species enters the community. *Proceedings of the National Academy of Sciences, 107,* 2015620163.

Greenfield, P. M. (1998). The cultural evolution of IQ. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 81–123). Washington, DC: American Psychological Association.

Hansen, J., & Wänke, M. (2010). Truth from language and truth from fit: the impact of linguistic concreteness and level of construal on subjective truth. *Personality and Social Psychology Bulletin, 36,* 1576–1588.

Hansen, M. T., & Haas, M. R. (2001). Competing for attention in knowledge markets: electronic document dissemination in a management consulting company. *Administrative Science Quarterly, 46,* 1–28.

Hawthorne, N. (2004). *The scarlet letter.* New York: Simon and Schuster.

Hills, T. T., & Adelman, J. S. (2015). Recent evolution of learnability in American English from 1800 to 2000. *Cognition*, 143, 87–92.

Hodgson, C., & Ellis, A. W. (1998). Last in, first to go: age of acquisition and naming in the elderly. *Brain and Language, 64,* 146–163.

Hollmann, W. B. (2009). Semantic change. In J. Culpeper, F. Katamba, P. Kerswill, & T. McEnery (Eds.), *English language: Description, variation and context* (pp. 525–537). Basingstoke: Palgrave.

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

Huang, H.-W., Lee, C.-L., & Federmeier, K. D. (2010). Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts. *Neuroimage, 49,* 1116–1123.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 130–136.
Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131,* 684–712.
Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolin gual word learning. *Psychonomic Bulletin & Review, 19,* 935–941.
Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General, 140,* 14–34.
Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44,* 978–990.
Labov, W. (1980). *Locating language in time and space.* New York: Academic Press.
Lieberman, E., Michel, J.-B., Jackson, J., Tang, T., & Nowak, M. A. (2007). Quantifying the evolutionary dynamics of language. *Nature, 449,* 713–716.
Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one, 5,* e8559.
Luther, D. (2009). The influence of the acoustic community on songs of birds in a neotropical rain forest. *Behavioral Ecology, 20,* 864–871.
McDaniel, M. A. (2006). Estimating state IQ: Measurement challenges and preliminary correlates. *Intelligence, 34,* 607–619.
MacKay, D. J. (2003). *Information theory, inference, and learning algorithms.* Cambridge: Cambridge University Press.
McWhorter, J. H. (2007). *Language interrupted: Signs of non-native acquisition in standard language grammars.* Oxford: Oxford University Press.
Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, ... Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science, 331,* 176–182.
Mickes, L., Darby, R. S., Hwe, V., Bajic, D., Warker, J. A., Harris, C. R., & Christenfeld, N. J. (2013). Major memory for microblogs. *Memory & Cognition, 41,* 481–489.
Miller, L. M., & Roodenrys, S. (2009). The interaction of word frequency and concreteness in immediate serial recall. *Memory & Cognition, 37,* 850–865.
Monaghan, P. (2014). Age of acquisition predicts rate of lexical evolution. *Cognition, 133,* 530–534.
Murphy, G. L., & Smith, E. E. (1982). Basic-level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior, 21,* 1–20.
Pagel, M., Atkinson, Q. D., & Meade, A. (2007). Frequency of word-use predicts rates of lexical evolution throughout indo-european history. *Nature, 449,* 717–720.
Paivio, A. (1971). *Imagery and verbal processes.* Hillsdale, New Jersey: Holt, Rinehart & Winston.
Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1196–1204.
Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76,* 1–25.
Pinker, S. (2014). *The sense of style: The thinking person's guide to writing in the 21st century.* London: Penguin.
Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 163178.

Rice, W. (1996). Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature, 381,* 232–234.

Romaine, S., Hogg, R. M., Blake, N. F., Lass, R., Algeo, J., & Burchfield, R. (1992). *The Cambridge history of the English language.* Cambridge: Cambridge University Press.

Romani, C., McAlpine, S., & Martin, R. C. (2008). Concreteness effects in different tasks: Implications for models of short-term memory. *The Quarterly Journal of Experimental Psychology, 61,* 292–323.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyesbraem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Sadoski, M. (2001). Resolving the effects of concreteness on interest, comprehension, and learning important ideas from text. *Educational Psychology Review, 13,* 263–281.

Sahami, M., Dumais, S., Heckerman, D., & Horvitz, E. (1998). A Bayesian approach to filtering junk e-mail. In *Learning for text categorization: papers from the 1998 workshop* (Vol. 62, pp. 98–105).

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27,* 499–520.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana, IL: University of Illinois Press.

Sidaris, D. (1994). *Barrel fever.* New York: Little, Brown and Company.

Simon, H. A. (1971). Designing organizations for an information-rich world. In M. Greenberge (Ed.), *Computers, communications, and the public interest* (pp. 37–52). Baltimore, MD: Johns Hopkins Press.

Tokowicz, N., & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes, 22,* 727–779.

Trudgill, P. (2002). *Sociolinguistic variation and change.* Baltimore, MD: Georgetown University Press.

Varian, H. R., & Lyman, P. (2000). How much information. *University of California, Berkeley School of Information Management & Systems Report: http://www.sims.berkeley.edu/how-much-info.*

Young, R. W., Morgan, W., & Midgette, S. (1992). *Analytical lexicon of Navajo.* Albuquerque, NM: University of New Mexico Press.

Zipf, G. K. (1949). *Human behavior and the principle of least effort.* Oxford: Addison-Wesley Press.