

**A Frequency Dictionary of  
Contemporary American English:  
Word Sketches, Collocates, and  
Thematic Lists**

# **Based On:**

THE CORPUS OF CONTEMPORARY  
AMERICAN ENGLISH (COCA)

400+ MILLION WORDS, 1990-2009

Created by

MARK DAVIES  
BRIGHAM YOUNG UNIVERSITY

[www.americancorpus.org](http://www.americancorpus.org)

- **Spoken: (79 million words)** Transcripts of unscripted conversation from more than 150 different TV and radio programs (examples: All Things Considered (NPR), Newshour (PBS), Good Morning America (ABC), Today Show (NBC), 60 Minutes (CBS), Hannity and Colmes (Fox), Jerry Springer, etc). [See notes on the naturalness and authenticity of the language from these transcripts).
- **Fiction: (76 million words)** Short stories and plays from literary magazines, children's magazines, popular magazines, first chapters of first edition books 1990-present, and movie scripts.
- **Popular Magazines: (81 million words)** Nearly 100 different magazines, with a good mix (overall, and by year) between specific domains (news, health, home and gardening, women, financial, religion, sports, etc). A few examples are Time, Men's Health, Good Housekeeping, Cosmopolitan, Fortune, Christian Century, Sports Illustrated, etc.
- **Newspapers: (76 million words)** Ten newspapers from across the US, including: USA Today, New York Times, Atlanta Journal Constitution, San Francisco Chronicle, etc. In most cases, there is a good mix between different sections of the newspaper, such as local news, opinion, sports, financial, etc.
- **Academic Journals: (76 million words)** Nearly 100 different peer-reviewed journals. These were selected to cover the entire range of the Library of Congress classification system (e.g. a certain percentage from B (philosophy, psychology, religion), D (world history), K (education), T (technology), etc.), both overall and by number of words per year

## 494 break v

n .law, heart, news, .rule, silence, story, .ground, .barrier, leg, bone, .piece, .neck, arm, .cycle, voice. misc .into, .away, .free, .apart, .loose

**up** marriage, .fight, boyfriend, meeting., girlfriend, union, band, pass, .demonstration, .monotony **down** .into, .barrier, car., .cry, .door, .tear, talk., enzyme., completely, negotiation. **out** war., fight., fire., sweat, fighting., riot., violence., .laugh, .hive **off** piece, talk, .engagement, negotiation, branch, abruptly, .relation

72917 | 0.97

### **3396 hypothesis *n***

*j* null, following, consistent, alternative, working, general, initial, original, theoretical, competing *n* study, support., result, test, research, testing, evidence, analysis, method, set *v*  
.*predict*, suggest, reject, examine, confirm, base, develop, formulate, *.state*, *.explain*

9282 | 0.82 A

“A” at the bottom indicates that the lemma *hypothesis* occurs at least twice as frequently in the Academic genre as it does overall in the corpus (Spoken, Fiction, Magazines, Newspapers).

# Actual Dictionary Entry

## 1598 lead *n*

adj early, commanding, female, huge, male, romantic, solid, molten, outstanding noun .role, .singer, level, .story, time, .paint, .poisoning, blood verb take., follow., hold., play, blow., build., contain., maintain.

21175 | 0.95

# Heterophones

1598 lead *n*

adj early, commanding, female, huge, male, romantic, solid, **molten**, outstanding noun .role, .singer, **level**, .story, time, **.paint**, **.poisoning**, **blood** verb take., follow., hold., play, blow., build., **contain.**, maintain.

21175 | 0.95

# Homographs

1598 lead *n*

adj early, commanding, female, huge, male, romantic, solid,  
molten, outstanding noun .role, .singer, level, .story,  
time, .paint, .poisoning, blood verb take., follow., hold., play,  
blow., build., contain., maintain.

21175 | 0.95



# Actual Dictionary Entry

## 4139 bow *n*

adj red, little, black, slight, deep, pink, short,

final noun .tie, .arrow, boat, shot., string, ship, .stern, hair,

line, wave verb take., draw, shoot, wear., pull., carry., bend,  
hang, .scrape, attach

6268 | 0.90 F M

# Heterophones

4139 bow *n*

adj red, little, black, slight, deep, pink, short,

final noun .tie, .arrow, boat, shot., string, ship, .stern, hair,  
line, wave verb take., draw, shoot, wear., pull., carry., bend,  
hang, .scrape, attach

6268 | 0.90 F M

[1 2009 FIC Bk:HusbandsWicked](#) Is there something I can do for you? " # He offered another **little bow**. “

[2 2009 FIC Bk:BoundSouth](#) brown Armani pants that I wear with a black silk shirt that ties in **little bows** around the cuffs.

[3 2009 FIC Bk:MortalCoils](#) The man nodded at Eliot and took a **little bow**.

[4 2009 FIC Bk:StrangerInNight](#) The man gave an awkward **little bow**.

[6 2008 FIC Bk:SecretService](#) lace on the high collar, bodice, back, and sleeves, with **little bows** on the skirt along with hundreds of tiny tucks.

[7 2008 FIC FantasySciFi](#) But sometimes, things don't tie up in a pretty **little bow** with appropriate theme music, Dazzle thought. Life just unravels until there's nothing

# Homographs

4139 bow *n*

adj red, little, black, slight, deep, pink, short,

final noun .tie, .arrow, boat, shot., string, ship, .stern, hair,

line, wave verb take., draw, shoot, wear., pull., carry., bend,

hang, .scrape, attach

6268 | 0.90 F M

### 198 run v

noun .president, business, .office, campaign, .risk, .finger, ad, .mile, .yard, race, candidate, operation, .Senate, ball, dog misc .into, .through, .away, home, .off, .across, hit., .fast, .deep, .wild, scream

**down** tear., .face, .cheek, .hall, blood., sweat., .spine, .list, clock, battery **out** time, money, luck., clock, patience, supply, benefit., .onto, limitation. **up** .down, .against, debt, .spine, .deficit, .huge, .score, .credit, rating. **around** .scream, .naked, .circle

172718 | 0.96

### 1283 run n

adj long, short, earned, winning, presidential, successful, dry, wild noun home., hit, .kick, game, touchdown., record, average, mile, salmon, bull. verb make., score, drive., allow., break., .fail, end, finish, earn

28805 | 0.92 N

## Animals

Note that several of these animals are also the mascot for sports teams or have figurative meaning (e.g. pig, mole), which would increase their overall frequency, and most of these are marked with parentheses in the following list.

[Top 80] dog n 49897, fish n 41277, bird 35610, horse 30042, chicken 23955, cat 20463, (bear) n 19980, fox 17303, turkey 15445, wolf 14452, deer 11259, (duck) n 11003, (tiger) 10210, cow 9755, mouse 9697, (eagle) 9395, snake 9387, lion 9260, rat 8115, (pig) 8048, (buffalo) 7678, cattle 7636, (hawk) n 7587, whale 6750, sheep 6172, bee 6148, shark 6035, rabbit 5927, monkey 5611, elephant 5562, goat 5330, worm n 5112, crab 4873, butterfly 4869, turtle 4667, crow n 4398, oyster 4328, frog 4304, goose n 4137, spider 4122, mosquito 4010, elk 3925, dolphin 3913, ant 3904, coyote 3850, lobster 3784, owl 3657, (falcon) 3626, mule 3111, (panther) 2972, (penguin) 2923, dove n 2906, squirrel 2832, camel 2825, (raven) 2673, (beetle) 2612, hog n 2298, moose 2231, pigeon 2225, ape 2203, pony 2194, swan 2059, donkey 1920, beaver 1895, (mole) 1860, gorilla 1694, alligator 1588, hare 1502, parrot 1456, crocodile 1426, bison 1327, leopard 1286, toad 1266, sparrow 1252, antelope 1168, quail 1122, ox 1117, raccoon 1092, gull 1068, heron 1057

## Clothing

### **General terms and parts of clothing**

(pocket) n 22655, (button) 13028, clothing 11590, (hood) 6800, jewelry 6052, sleeve 5607, collar 5564, cuff 3276

### **Items of clothing**

(suit) n 35280, shoe 26007, (ring) n 21069, shirt 20470, dress n 20389, hat 17462, (tie) n 17060, coat n 15297, jacket 14938, boot n 14436, belt n 13966, pants 11412, glove 9852, uniform n 9809, skirt n 8858, jeans 8037, sock n 6486, sweater 5179, robe 4751, shorts 4626, gown 4387, scarf n 3551, (slip) n 3545, vest 3313, blouse 3110, underwear 3037, necklace 2914, diaper 2738, earring 2720, cloak 2624, bracelet 2394, bra 2378, apron 2205, sneakers 2181, stocking 2027, slipper 1943, blazer 1932, pajamas 1491, bikini 1418, sweatshirt 1342, baseball cap 1087, nightgown 1031, high heels 985, overcoat 979, parka 859, overalls 812, swimsuit 799, tights 729, (loafers) 677, raincoat 670, cardigan 611, bow tie 593, halter 591, mitten 583, cowboy boot 561, tank top 553, underpants 489, leggings 442, undershirt 413, suspenders 384, polo shirt 369, petticoat 364, pantyhose 329, trousers 283, pullover 271, leotards 254, camisole 164, galosh(es) 111, sweatsuit 93, earmuffs 89

## **Fiction**

Lemmas that are at least three times as common (per million words) in fiction as in the other four major genres.

**[nouns]** hand 114904, eye 103343, head 80920, room 79276, door 77861, face 74067, mother 73338, father 61551, voice 48768, girl 46151, boy 45513, arm 44069, hair 39458, light 38758, window 37514, bed 32997, wall 30797, floor 28759, finger 25502, shoulder 25386

**[verbs]** look 206546, sit 70921, stand 65788, walk 55092, pull 45033, wait 44021, smile 31790, stare 31593, shake 29573, laugh 28586, close 25973, nod 23062, wonder 21782, hang 20956, lay 19814, lean 18584, notice 17968, step 17506, sleep 17349, touch 16890

**[adjectives]** dead 24327, dark 21706, sorry 18709, empty 12890, tall 11777, quiet 11446, strange 10645, thin 10006, alone 9897, thick 9705, silent 9438, gray 9427, brown 9427, yellow 9388, tired 8004, pale 7454, asleep 6837, glad 6523, stupid 6505, wet 6492

**[adverbs]** away 60710, maybe 45211, suddenly 23941, slowly 17663, inside 15466, please 15106, okay 13662, anyway 13263, outside 10418, somewhere 9846, straight 9730, barely 8109, closer 7400, quietly 7152, gently 6185, softly 5710, besides 4697, upstairs 3859, o'clock 3776, silently 3542



## **Academic Journals**

Lemmas that are at least three times as common (per million words) in academic journals as in the other four major genres

**[nouns]** student 163046, study 108068, teacher 73873, education 67286, level 61510, research 61043, community 56722, result 55682, process 53533, development 51798, use 51365, policy 48208, data 48070, effect 45702, experience 44830, activity 44133, model 42673, analysis 40610, behavior 40419, difference 40129

**[verbs]** provide 73643, suggest 40118, develop 39775, require 37581, base 33053, indicate 31471, describe 30752, identify 29310, represent 26974, increase 26032, present 24867, note 24542, determine 24423, occur 23936, relate 23803, establish 22807, examine 21117, state 20879, compare 20499, reflect 19508

**[adjectives]** social 80721, political 63690, economic 43584, significant 37691, cultural 30243, environmental 29711, physical 27783, specific 23780, similar 23717, individual 23127, various 21850, religious 21673, positive 20001, traditional 19861, academic 19351, African 19323, sexual 19305, particular 18709, present 18065, effective 17667

**[adverbs]** however 70754, thus 39524, (for) example 36060, therefore 21538, e.g. 18417, significantly 15657, generally 14601, highly 12450, (in) addition 12087, relatively 12025, i.e. 11234, moreover 10350, frequently 9231, specifically 9077, primarily 8981, approximately 8787, furthermore 7452, similarly 7159, previously 6989, effectively 6415

## New words in American English

Lemmas that are at least three times as common (per million words) in the texts from the 2000s (2000-2008; 180+ million words) as in the 200+ million words from the 1990s.

**[nouns]** e-mail 14326, terrorism 10360, terrorist 8889, affiliation 8713, adolescent 7212, homeland 4157, website 3909, Sunni 3637, wireless 3492, prep 3362, Taliban 3006, insurgent 2433, globalization 2354, SUV 2068, RPG 1970, anthrax 1954, steroid 1898, genome 1867, blog 1765, detainee 1733, militant 1610, ethanol 1601, insurgency 1561, yoga 1533, recount 1491, cleric 1466, coping 1380, tsunami 1310, cellphone 1149

**[verbs]** host 5535, click 4094, e-mail 2139, download 1851, preheat 1364, bully 916, makeover 853, freak 649, partner 638, mentor 587, morph 349, vaccinate 286, restart 258, reconnect 257, sauté 251, hijack 245, co-write 236, ditch 183, reference 148, swipe 142, outsource 141, transition 136, upload 127, refuel 126, profile 122, encrypt 117, workout 117, prep 115, splurge 114, snack 106

**[adjectives]** online 9219, terrorist 7908, Afghan 2776, Taliban 2095, Shiite 2052, Pakistani 1967, same-sex 1397, sectarian 1138, upscale 1057, embryonic 1036, Islamist 929, iconic 765, faith-based 754, broadband 738, handheld 734, pandemic 719, web-based 711, nonstick 689, steroid 655, insurgent 586, avian 553, dot-com 532, Chechen 396, old-school 360, clueless 356, performance-enhancing 350, high-stakes 337, Al-Qaida 331, 21st-century 331, gated 326

**[adverbs]** online 6034, famously 1173, postoperatively 226, offline 158, wirelessly 141, healthfully 72, preemptively 66, intraoperatively 58, triply 33, day-ahead 30, forensically 28, inferiorly 26, pre-emptively 25, multiculturally 22, counterintuitively 21, synchronically 21

## +American (COCA) –British(BNC)

The following are just a sampling of the words that are much more common in COCA than in the BNC. In all cases, the word is at least twice as far down the list in the BNC wordlist as in COCA (e.g. #2000 in COCA, #5000 in BNC).

**Verbs:** call, report, focus, guess, sign, step, figure, roll, fire, hire, file, oppose, wrap, interview, accomplish, testify, bake, track, evolve, violate, target, pitch, flip, ruin, hike, invade

**Nouns:** student, president, percent, kid, guy, nation, photo, arm, American, Republican, phone, movie, store, lawyer, Democrat, professor, expert, senator, break, camera, coach, item, mom, dream, attorney, scientist, web, camp, truck, apartment, bowl, baseball, internet, basketball

**Adjectives:** American, federal, tough, native, Iraqi, crazy, smart, Israeli, Mexican, congressional, elementary, online, gifted, athletic, ongoing, African-american, suburban, Hispanic, scary, high-tech, cute, nonprofit, immigrant, skeptical, aging, low-income, interstate

## +British (BNC) / -American (COCA)

The following are just a sampling of the words that are much more common in the BNC than in COCA. In all cases, the word is at least twice as far down the list in the COCA wordlist as in the BNC (e.g. #2000 in the BNC, #5000 in COCA).

**Verbs:** ensure, suppose, regard, voice, bind, retain, undertake, phone, wan, allocate, knit, book, abolish, envisage, incur, fancy, commence, enclose, enquire, sack, adjourn, tidy, query, retort, queue, nick, remand, smelt

**Nouns:** council, minister, union, pound, scheme, shop, principle, village, provision, sector, appeal, parliament, mum, tea, lord, cabinet, pension, flat, expenditure, solicitor, coal, crown, castle, pub, parish, tenant, councillor, countryside, plaintiff, autumn, cottage

**Adjectives:** British, European, English, royal, French, industrial, Scottish, lovely, working, bloody, parliamentary, alright, statutory, keen, Welsh, Tory, socialist, unemployed, delighted, Victorian, liable, superb, working-class, compulsory, splendid, post-war, dreadful, redundant, inland, wee

**A Frequency Dictionary of  
Contemporary American English:  
Word Sketches, Collocates, and  
Thematic Lists**

Mark\_Davies@byu.edu

Dee\_Gardner@byu.edu