

Towards Automatic Finding of Word Sense Changes in Time

Vít Baisa^{1,2}, Ondřej Herman¹, Miloš Jakubíček^{1,2}

¹ Natural Language Processing Centre
Faculty of Informatics, Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic
{xherman1,xbaisa,jak}@fi.muni.cz

² Lexical Computing
Brighton, United Kingdom and Brno, Czech Republic
{vit.baisa,milos.jakubicek}@sketchengine.co.uk

Abstract. We present a methodology proposal for finding changes in contextual behaviour of words. Assuming that a word sense is defined by actual usages of the word in a given context, this task corresponds to finding changes of word senses. We outline here main ideas of our distributional approach based on word sketches and discuss preliminary results.

Keywords: neologism, Sketch Engine, word sense, word sketch

1 Introduction

Thanks to technology advances, linguistics has undergone a major paradigm shift over the past 20 years—the field, previously largely dominated by an introspection, can now benefit from a vast amount of empirical evidence available in text corpora, incepting a new subfield—corpus linguistics. Most of the text corpora are synchronous—they represent a snapshot of some specific text type or general language at some time point. As such, they are suitable for many kinds of a synchronic analysis.

A diachronic analysis, i.e. tracking language changes over time, requires diachronic corpora carefully collected over large time spans, ideally with very balanced content (the same text types and text size from each time span). For English, some of the available corpora include the BNC [1] (which is however quite outdated and individual time spans are not very balanced), the COCA [2], the OEC [3] or the Feeds corpus [4]. A valuable resource, though not a text corpus, is also the Google Books n-grams database [5].

In this paper we discuss an extension to the existing trends finding functionality within Sketch Engine, a leading corpus management tool [6]. Sketch Engine has been recently [4] enhanced with a diachronic analysis function that tracks lexical changes in text corpora and estimates trends for individual words. This is useful also for neologism finding—however, many neologisms are not

new word forms (lexemes) but new word senses. Building on the assumption that a word's sense is defined by the context in which the word is used, we propose a new approach (based on word sketches) for tracking changes of the context of a word which might often indicate a shift in its sense.

2 Sketch Engine

Sketch Engine is an online corpus management system providing access to hundreds of text corpora which can be searched and analyzed. It gained its name after one of its key features—word sketches, one page summaries of a word's collocational behaviour, in particular grammatical relations (see Figure 1). In technical terms, word sketches represent dependency syntax with each table item being a dependency triple consisting of a headword, a relation and a collocation, such as *resource–modifier–scarce*. These triples are scored using a lexicographic association score, in this case the logDice [7].

resource (*noun*) British National Corpus freq = [12658](#) (112.8 per million)

<u>modifier</u>	6477	1.5	<u>object of</u>	3285	2.2	<u>modifies</u>	1906	0.5	<u>subject of</u>	512	0.6
scarce	163	9.53	allocate	194	9.58	allocation	135	9.42	devote	28	7.69
natural	321	8.94	pool	39	8.43	implication	46	7.09	consume	4	5.36
limited	187	8.86	exploit	64	8.23	management	153	6.98	tie	6	4.87
financial	249	8.3	divert	38	7.86	defense	7	6.68	last	4	4.6
mineral	89	8.19	deploy	31	7.67	Stonier	6	6.65	back	5	4.5
additional	107	7.92	devote	44	7.64	utilisation	7	6.63	stretch	4	4.29
valuable	74	7.86	concentrate	62	7.35	committee	132	6.49	result	6	3.93
extra	88	7.53	utilise	22	7.28	centre	158	6.4	depend	6	3.84
human	134	7.38	conserve	17	7.09	allocator	5	6.4	limit	5	3.59
renewable	33	7.31	lack	37	7.0	depletion	6	6.21	match	3	3.58
adequate	49	7.28	reallocate	13	6.98	pack	17	6.2	share	6	3.55
non renewable	25	6.97	mobilise	13	6.83	investigator	8	6.17	earn	3	3.55
existing	53	6.68	mobilize	13	6.79	column	20	6.16	enable	7	3.54
finite	22	6.66	distribute	29	6.73	constraint	14	6.14	remain	12	3.5

Fig. 1. An example of a word sketch for the noun *resource*.

Provided there is diachronic annotation available in a corpus, the system can be setup to offer a trends finding feature. This operates on individual corpus tokens with associated positional attributes such as word form, lemma or part-of-speech tag. The frequency distribution of the input attribute value is subject to regression analysis (linear regression or Theil-Sen estimator) that estimates a trend defined by a curve slope. The user is then presented with an ordered list of most (positively or negatively) trending words as given in Figure 2.

word	Trend	p-value	Freq	Graph
<u>extremists</u>	3.7320 +	0.002108	<u>2,597</u>	
<u>unbelievers</u>	3.7320 +	0.003182	<u>331</u>	
<u>contraband</u>	3.7320 +	0.006029	<u>541</u>	
<u>stoppages</u>	3.7320 +	0.006029	<u>266</u>	
<u>carves</u>	3.7320 +	0.003182	<u>179</u>	
<u>three-and-a-half</u>	3.7320 +	0.006029	<u>417</u>	
<u>cellist</u>	3.7320 +	0.003182	<u>384</u>	
<u>newly-created</u>	3.7320 +	0.003182	<u>130</u>	
<u>deadpan</u>	3.7320 +	0.003182	<u>337</u>	
<u>auteur</u>	3.7320 +	0.003182	<u>248</u>	
<u>whitewashing</u>	3.7320 +	0.003182	<u>135</u>	
<u>trapeze</u>	3.7320 +	0.003182	<u>184</u>	
<u>outstripping</u>	3.7320 +	0.003182	<u>250</u>	
<u>analytically</u>	3.7320 +	0.003182	<u>135</u>	
<u>people-to-people</u>	3.7320 +	0.003182	<u>114</u>	
<u>nonpartisan</u>	3.7320 +	0.003182	<u>672</u>	
<u>rollicking</u>	3.7320 +	0.003182	<u>279</u>	
<u>masonry</u>	3.7320 +	0.003182	<u>391</u>	
<u>shareable</u>	3.7320 +	0.003182	<u>415</u>	
<u>transgender</u>	3.7320 +	0.003182	<u>2,719</u>	
<u>devolving</u>	3.7320 +	0.003182	<u>174</u>	
<u>takeaways</u>	3.7320 +	0.003182	<u>1,034</u>	
<u>wags</u>	3.7320 +	0.003182	<u>129</u>	
<u>intraday</u>	3.7320 +	0.003182	<u>666</u>	
<u>bleakest</u>	3.7320 +	0.003182	<u>89</u>	

Fig. 2. Most trending word forms in the Feeds corpus computed using the Theil-Sen estimation.

3 Finding Word Sense Changes

Sketch Engine identifies the most salient contexts in the word sketch feature. Therefore, we decided to experiment with using the existing trends computation on the top of word sketches. The input values for the regression analysis were frequency counts of the word sketch triples in the individual time spans. This way we obtained most trending triples which can be subject to further analysis (e.g. clustering of headwords). In Table 1 we list some linguistically interesting items found in the top 100 triples (both increasing and decreasing trends, sorted by the slope of the regression curve) identified in the COCA which covers a twenty-year period from 1990 until 2010, having 18 million tokens for each year (and additional 14 and 8 millions for 2011 and 2012).

Out of the top 100 increasing and top 100 decreasing triples only 19 are not noun-noun or adjective-noun modifiers. Clearly taking such triple list without further modification yields mainly trending multi-word noun phrases but not so many verb patterns. Therefore we also experimented with grouping the triples by headword. For each headword we calculated its score as the

Table 1. Sample trending triples (positive and negative) computed in the COCA.

+	–
peat-n modifies forest-n	percentage-n modifies plan-n
social-j modifies gradient-n	new-j modifies atheist-n
mastery-n modifies climate-n	diadromous-j modifies fish-n
joint-j modifies household-n	rate-n pp_of maltreatment-n
transaction-n modifies approach-n	medium-n modifier dual-j
trauma-n modifies theory-n	description-n pp_obj_for microfilm-n
learn-v object common-n	talk-v object pedometer-n
sexual-j and/or subjective-j	reverse-j modifies student-n
middling-j modifies sort-n	common-j modifies carriage-n
adult-j modifies day-n	new-j and/or packed-j
chick-n modifies lit-n	downwind-j modifies state-n
diminish-v subject regulation-n	cell-n object_of reprogram-v

multiplication of absolute values of all slopes in triples of this headword (which corresponds to composing the regression curves). Since the slope is often a small floating point number, we operate in an additive logarithmic calculus instead of directly multiplicative one.

The resulting scoring was clustered by part-of-speech of the headword into 4 groups: nouns, adjectives, verbs and others. In Tables 2–5 we provide the most changed headwords obtained together with the top 5 changed collocates—both positive (\oplus) and negative (\ominus):

4 Conclusions

We have presented an early stage research on novel methods for corpus-based finding of changes in word senses. The achieved results are promising but clearly indicate that more effort must be put into clustering of the word sketch triples and providing more insight into corpus data that would allow an easy interpretation of the results.

Acknowledgements This work has been partly supported by the Ministry of Education of CR within the Czech-Norwegian Research Programme in the HaBiT Project 7F14047.

References

1. Leech, G.: 100 million words of English: the British National Corpus (BNC). *Language Research* **28**(1) (1992) 1–13

2. Davies, M.: The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and linguistic computing* (2010)
3. Press, O.U.: *Oxford english corpus* (2015)
4. Kilgarriff, A., Herman, O., Bušta, J., Rychlý, P., Jakubíček, M.: DIACRAN: a framework for diachronic analysis. (2015)
5. Lin, Y., Michel, J.B., Aiden, E.L., Orwant, J., Brockman, W., Petrov, S.: Syntactic annotations for the google books ngram corpus. In: *Proceedings of the ACL 2012 System Demonstrations*. ACL '12, Stroudsburg, PA, USA, Association for Computational Linguistics (2012) 169–174
6. Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., Suchomel, V.: The Sketch Engine: ten years on. *Lexicography* **1** (2014)
7. Rychlý, P.: A lexicographer-friendly association score. *Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN* (2008) 6–9

Table 2. Most trending nouns in COCA computed by grouping word sketch triples.

time-n	⊖constant-j, ⊖post-n, ⊖bat-n, ⊖unit-n, ⊕rating-n, ⊕hour-n, ⊕working-j, ⊕address-n
people-n	⊖distinguish-v, ⊖street-n, ⊖welfare-n, ⊖boat-n, ⊖wish-n, ⊕ensure-v, ⊕transgender-n, ⊕amazing-j, ⊕other-n, ⊕autism-n
year-n	⊖company-n, ⊖communism-n, ⊖recession-n, ⊖crucial-j, ⊕gender-n, ⊕collection-n, ⊕pound-n, ⊕multiple-j, ⊕old-j
man-n	⊖indigenous-j, ⊖advance-n, ⊖woman-n, ⊖public-j, ⊕slave-n, ⊕present-v, ⊕detain-v, ⊕connect-v
woman-n	⊖old-n, ⊖reasonable-j, ⊖implant-n, ⊖black-n, ⊖man-n, ⊕integration-n, ⊕planet-n, ⊕focus-v, ⊕acceptance-n, ⊕use-n
system-n	⊖investment-n, ⊖ledger-n, ⊖interpretation-n, ⊖ductal-j, ⊖bus-n, ⊕assurance-n, ⊕da-n, ⊕variable-n, ⊕multiprocessor-n, ⊕aquifer-n
group-n	⊖whale-n, ⊖nominal-j, ⊖genetic-j, ⊖soil-n, ⊖correlation-n, ⊕supervision-n, ⊕substance-n, ⊕trained-j, ⊕delayed-j, ⊕exposure-n
child-n	⊖transfer-n, ⊖diabetic-j, ⊖temperament-n, ⊖marriage-n, ⊖creative-j, ⊕implant-n, ⊕underachieve-v, ⊕injury-n, ⊕difficulty-n, ⊕rape-n
life-n	⊖tribal-j, ⊖answer-n, ⊖style-n, ⊖element-n, ⊖change-n, ⊕relationship-n, ⊕coach-n, ⊕youth-n, ⊕battery-n, ⊕nomadic-j
student-n	⊖reverse-j, ⊖handicap-n, ⊖treatment-n, ⊖retarded-j, ⊖nondisabled-j, ⊕voucher-n, ⊕undocumented-j, ⊕nurse-v, ⊕algebra-n, ⊕bully-v
program-n	⊕supervisory-j, ⊖timber-n, ⊖thinking-n, ⊖tune-v, ⊖export-n, ⊕bully-v, ⊕yard-n, ⊕trimming-n, ⊕meditation-n, ⊕rating-n
work-n	⊖comparative-j, ⊖record-v, ⊖video-j, ⊖art-n, ⊖wage-n, ⊕curriculum-n, ⊕profession-n, ⊕travel-v, ⊕theory-n, ⊕literature-n
state-n	⊖downwind-j, ⊖store-n, ⊖terror-n, ⊖red-j, ⊖density-n, ⊕confidence-n, ⊕space-n, ⊕upwind-j, ⊕failed-j, ⊕pool-n
school-n	⊖victimization-n, ⊖field-n, ⊖samba-n, ⊖bureaucracy-n, ⊖entry-n, ⊕bully-v, ⊕bully-v, ⊕intervention-n, ⊕liaison-n, ⊕characteristic-n
way-n	⊖awkward-j, ⊖sensible-j, ⊖base-n, ⊖people-n, ⊕fun-n, ⊕point-n, ⊕big-j, ⊕food-n, ⊕indigenous-j
number-n	⊖seal-n, ⊖fisherman-n, ⊖stock-n, ⊖share-n, ⊖black-n, ⊕accounting-n, ⊕hometown-n, ⊕skill-n, ⊕cumulative-j, ⊕cell-n
day-n	⊖set-n, ⊖strike-n, ⊖government-n, ⊖quote-n, ⊖collection-n, ⊕adult-j, ⊕step-n, ⊕lust-n, ⊕smoke-v, ⊕be-v
area-n	⊖nursery-n, ⊖writer-n, ⊖cooperation-n, ⊖goal-n, ⊖development-n, ⊕story-n, ⊕wellhead-n, ⊕thought-n, ⊕necrosis-n, ⊕receive-v
level-n	⊖ethic-n, ⊖signal-n, ⊖fandom-n, ⊖cognition-n, ⊖output-n, ⊕identification-n, ⊕creatinine-n, ⊕lipid-n, ⊕parent-n, ⊕achievement-n

Table 3. Most trending adjectives in COCA computed by grouping word sketch triples.

more-j	⊖stock-n, ⊖hill-n, ⊖reward-n, ⊖belief-n, ⊕insulin-n, ⊕aggression-n, ⊕transparency-n, ⊕fantasy-n, ⊕green-n
such-j	⊖industry-n, ⊖special-j, ⊖good-n, ⊖stock-n, ⊖conduct-n, ⊕websites-n, ⊕library-n, ⊕meat-n, ⊕flavor-n, ⊕nurse-n
other-j	⊖saurs-n, ⊖match-n, ⊖fishery-n, ⊖payer-n, ⊕troll-n, ⊕covariates-n, ⊕websites-n, ⊕scandal-n, ⊕node-n
new-j	⊖atheist-n, ⊖packed-j, ⊖commonwealth-n, ⊖landfill-n, ⊖powder-n, ⊕physicist-n, ⊕normal-n, ⊕website-n, ⊕subscription-n, ⊕normal-j
good-j	⊖catfish-n, ⊖crappie-n, ⊖opponent-n, ⊖dealer-n, ⊕competition-n, ⊕teammate-n, ⊕bass-n, ⊕climber-n, ⊕learning-n
many-j	⊖skier-n, ⊖peasant-n, ⊖composer-n, ⊖debate-n, ⊖ingredient-n, ⊕intervention-n, ⊕sample-n, ⊕trend-n, ⊕challenge-n, ⊕challenge-n
first-j	⊖black-n, ⊖offender-n, ⊖complaint-n, ⊖instant-n, ⊕branchial-j, ⊕variate-n, ⊕responder-n, ⊕purpose-n, ⊕putt-n
own-j	⊖aide-n, ⊖superiority-n, ⊖prediction-n, ⊖century-n, ⊖psyche-n, ⊕drug-n, ⊕bar-n, ⊕take-n, ⊕online-j, ⊕wine-n
small-j	⊖orchard-n, ⊖stock-n, ⊖have-v, ⊖investor-n, ⊕channel-n, ⊕reactor-n, ⊕frequent-j, ⊕robot-n, ⊕buck-n
great-j	⊖bulk-n, ⊖dependence-n, ⊖heap-n, ⊖famine-n, ⊖loss-n, ⊕recession-n, ⊕face-n, ⊕mom-n, ⊕option-n, ⊕cast-n
little-j	⊖wing-n, ⊖dragon-n, ⊖look-n, ⊖argument-n, ⊕shabby-j, ⊕speak-v, ⊕camera-n, ⊕close-j, ⊕egg-n
large-j	⊖republic-n, ⊖laboratory-n, ⊖customer-n, ⊖lobster-n, ⊖hotel-n, ⊕heat-n, ⊕raw-j, ⊕lemon-n, ⊕nonstick-j, ⊕oven-n
high-j	⊖subject-n, ⊖scenario-n, ⊖rent-n, ⊖stage-n, ⊕gifted-j, ⊕youth-n, ⊕odd-n, ⊕dive-n, ⊕definition-n, ⊕career-n
old-j	⊖dance-n, ⊖economic-j, ⊖communist-j, ⊖viewer-n, ⊕year-n, ⊕participant-n, ⊕report-n, ⊕cartoon-n, ⊕dumb-j
big-j	⊖cigar-n, ⊖holding-n, ⊖look-n, ⊖newspaper-n, ⊖gamble-n, ⊕physical-j, ⊕celebrity-n, ⊕presence-n, ⊕opening-n, ⊕army-n
few-j	⊖stock-n, ⊖supporter-n, ⊖black-n, ⊖department-n, ⊕tweak-n, ⊕click-n, ⊕emission-n, ⊕string-n, ⊕bone-n
most-j	⊖price-n, ⊖skier-n, ⊖operator-n, ⊖tribe-n, ⊖therapist-n, ⊕component-n, ⊕phone-n, ⊕participant-n, ⊕liquid-n, ⊕county-n
same-j	⊖nation-n, ⊖tendency-n, ⊖shoe-n, ⊖department-n, ⊖restriction-n, ⊕interview-n, ⊕skillet-n, ⊕scrutiny-n, ⊕target-n, ⊕academic-j
different-j	⊖filter-n, ⊖relation-n, ⊖society-n, ⊖movement-n, ⊖political-j, ⊕landscape-n, ⊕stuff-n, ⊕night-n, ⊕demographic-j, ⊕rate-n

Table 4. Most trending verbs in COCA computed by grouping word sketch triples.

be-v	⊖logos-n, ⊖pellagra-n, ⊖shamanism-n, ⊖digester-n, ⊕ingrethents-n, ⊕authence-n, ⊕texting-v, ⊕diat-n, ⊕hed-n
have-v	⊖anthrax-n, ⊖mage-n, ⊖chickenpox-n, ⊕ames-n, ⊕blog-n, ⊕website-n, ⊕antidepressant-n, ⊕memorial-n
make-v	⊖newspaper-n, ⊖subject-n, ⊖district-n, ⊖antibody-n, ⊕sidebar-n, ⊕resilient-j, ⊕minute-n, ⊕tweak-n, ⊕backswing-n
do-v	⊖opponent-n, ⊖trail-n, ⊖aid-n, ⊖black-n, ⊕sidebar-n, ⊕adolescent-n, ⊕librarian-n, ⊕diligence-n, ⊕golfer-n
use-v	⊖order-v, ⊖ecstasy-n, ⊖fastball-n, ⊖airline-n, ⊕ingrethents-n, ⊕braille-n, ⊕metric-n, ⊕cellphones-n, ⊕website-n
get-v	⊖rocket-n, ⊖position-n, ⊖welfare-n, ⊖hostage-n, ⊕sidebar-n, ⊕text-n, ⊕stop-n, ⊕overwhelmed-j, ⊕com-n
go-v	⊖com-n, ⊖pond-n, ⊖station-n, ⊕viral-j, ⊕tobazaar-n, ⊕prevention-n, ⊕voice-v, ⊕sfgate-n
take-v	⊖leadership-n, ⊖overdose-n, ⊖capsule-n, ⊖paper-n, ⊕statins-n, ⊕meds-n, ⊕phone-n, ⊕sidebar-n, ⊕assessment-n
say-v	⊖corpse-n, ⊖republic-n, ⊕jazz-n, ⊕ing-n, ⊕website-n, ⊕demon-n, ⊕release-n, ⊕utility-n, ⊕down-x
come-v	⊖opposition-n, ⊖occupy-v, ⊖tell-v, ⊖program-n, ⊕kitchen-n, ⊕ghost-n, ⊕loss-n, ⊕online-j, ⊕remark-n
see-v	⊖earning-n, ⊖contract-n, ⊖doll-n, ⊖section-n, ⊕hardcopy-n, ⊕website-n, ⊕reference-n, ⊕online-j, ⊕http-n
become-v	⊖expression-n, ⊖cliche-n, ⊖card-n, ⊖officer-n, ⊖scholar-n, ⊕submission-n, ⊕guy-n, ⊕brand-n, ⊕cell-n, ⊕distraction-n
include-v	⊖survivor-n, ⊖wife-n, ⊖stretch-n, ⊖son-n, ⊖fare-n, ⊕covariates-n, ⊕band-n, ⊕determine-v, ⊕win-n, ⊕awareness-n
find-v	⊖bass-n, ⊖comet-n, ⊖residue-n, ⊖variety-n, ⊕sidebar-n, ⊕lesion-n, ⊕online-j, ⊕online-j, ⊕improve-v
begin-v	⊖tale-n, ⊖strike-n, ⊖black-n, ⊖therefore-a, ⊖listen-v, ⊕split-v, ⊕edge-n, ⊕archaeologist-n, ⊕arrive-v, ⊕dough-n
look-v	⊖age-n, ⊖hank-n, ⊖growth-n, ⊖court-n, ⊕amazing-j, ⊕jack-n, ⊕pair-n, ⊕photo-n, ⊕character-n
give-v	⊖black-n, ⊖resistance-n, ⊕judgment-n, ⊕urgency-n, ⊕how-x, ⊕prevalence-n, ⊕foot-n, ⊕certificate-n, ⊕scientist-n
seem-v	⊖spring-v, ⊖agency-n, ⊖century-n, ⊖dubious-j, ⊕solution-n, ⊕fun-n, ⊕die-v, ⊕set-v, ⊕smart-j
work-v	⊖piece-n, ⊖location-n, ⊖deal-n, ⊖medium-n, ⊖class-n, ⊕move-n, ⊕participant-n, ⊕batch-n, ⊕capacity-n, ⊕fluid-n
need-v	⊖telescope-n, ⊖finance-v, ⊖athlete-n, ⊖defense-n, ⊖satisfy-v, ⊕practitioner-n, ⊕nurse-n, ⊕miracle-n, ⊕option-n, ⊕dig-v

Table 5. Most trending other words in COCA computed by grouping word sketch triples.

his-d	⊖phaser-n, ⊖deer-n, ⊖quilt-n, ⊖envelope-n, ⊕stunner-n, ⊕blog-n, ⊕falcon-n, ⊕website-n, ⊕email-n
their-d	⊖blogs-n, ⊖playoff-n, ⊖national-n, ⊖banker-n, ⊕websites-n, ⊕cellphones-n, ⊕episode-n, ⊕primary-n, ⊕website-n
her-d	⊖administration-n, ⊖prejudice-n, ⊖cross-n, ⊖abortion-n, ⊕blog-n, ⊕fianc-n, ⊕rsum-n, ⊕website-n, ⊕reticule-n
its-d	⊖quarry-n, ⊖prospectus-n, ⊖recession-n, ⊖printing-n, ⊕website-n, ⊕footprint-n, ⊕heft-n, ⊕electricity-n, ⊕win-n
my-d	⊖stepmom-n, ⊖ski-n, ⊖cellphone-n, ⊖speculation-n, ⊕integrator-n, ⊕blog-n, ⊕participant-n, ⊕website-n, ⊕fianc-n
your-d	⊖infant-n, ⊖wreath-n, ⊖ski-n, ⊖skate-n, ⊕smartphone-n, ⊕protagonist-n, ⊕gadget-n, ⊕aspect-n, ⊕downswing-n
not-a	⊖transcribe-v, ⊖inconsistent-j, ⊖persuasive-j, ⊖wise-j, ⊖recur-v, ⊕access-v, ⊕post-v, ⊕manipulate-v, ⊕predetermine-v, ⊕assess-v
our-d	⊖mama-n, ⊖century-n, ⊖intimacy-n, ⊖cave-n, ⊕website-n, ⊕forum-n, ⊕specialist-n, ⊕shoot-n, ⊕blog-n
also-a	⊖smile-v, ⊖judge-v, ⊖guarantee-v, ⊖concede-v, ⊖desire-v, ⊕additional-j, ⊕detail-v, ⊕inject-v, ⊕picture-v, ⊕s-v
so-a	⊖label-v, ⊖sanguine-j, ⊖gracious-j, ⊖destructive-j, ⊖broad-j, ⊕freak-v, ⊕base-v, ⊕improbable-j, ⊕worth-j, ⊕welcome-v
then-a	⊖react-v, ⊖mail-v, ⊖negotiate-v, ⊖record-v, ⊖plunge-v, ⊕chill-v, ⊕top-j, ⊕halt-v, ⊕post-v, ⊕remind-v
as-a	⊖cry-v, ⊖blank-j, ⊖radical-j, ⊖give-v, ⊖dependent-j, ⊕delicious-j, ⊕test-v, ⊕die-v, ⊕place-v, ⊕connect-v
only-a	⊖attach-v, ⊖award-v, ⊖alternative-j, ⊖legal-j, ⊖token-j, ⊕straight-j, ⊕embody-v, ⊕clean-v, ⊕organic-j, ⊕motivate-v
still-a	⊖fume-v, ⊖uncertain-j, ⊖viable-j, ⊖evoke-v, ⊖suspect-v, ⊕link-v, ⊕relevant-j, ⊕sport-v, ⊕throw-v, ⊕crave-v
just-a	⊖flat-j, ⊖issue-v, ⊖eliminate-v, ⊖cite-v, ⊖wan-v, ⊕combine-v, ⊕wilt-v, ⊕shy-v, ⊕focus-v, ⊕tap-v
even-a	⊖strict-j, ⊖strike-v, ⊖such-j, ⊖contradictory-j, ⊖deadly-j, ⊕tear-v, ⊕accelerate-v, ⊕result-v, ⊕happy-j, ⊕par-j
very-a	⊖frighten-v, ⊖rapid-j, ⊖restrictive-j, ⊖definite-j, ⊖minor-j, ⊕persistent-j, ⊕tender-j, ⊕crowded-j, ⊕thankful-j, ⊕experience-v
now-a	⊖own-j, ⊖handle-v, ⊖concentrate-v, ⊖yield-v, ⊖pend-v, ⊕block-v, ⊕avoid-v, ⊕date-v, ⊕host-v, ⊕easy-j
often-a	⊖succeed-v, ⊖less-j, ⊖disappear-v, ⊖force-v, ⊖similar-j, ⊕frame-v, ⊕opt-v, ⊕whisk-v, ⊕host-v, ⊕stir-v
never-a	⊖permit-v, ⊖begin-v, ⊖practice-v, ⊖succeed-v, ⊖press-v, ⊕shy-v, ⊕write-v, ⊕recommend-v, ⊕underestimate-v, ⊕late-j