

LING/C SC 581:

Advanced Computational Linguistics

Lecture 20

Prof. Sandiway Fong

Today's Topic

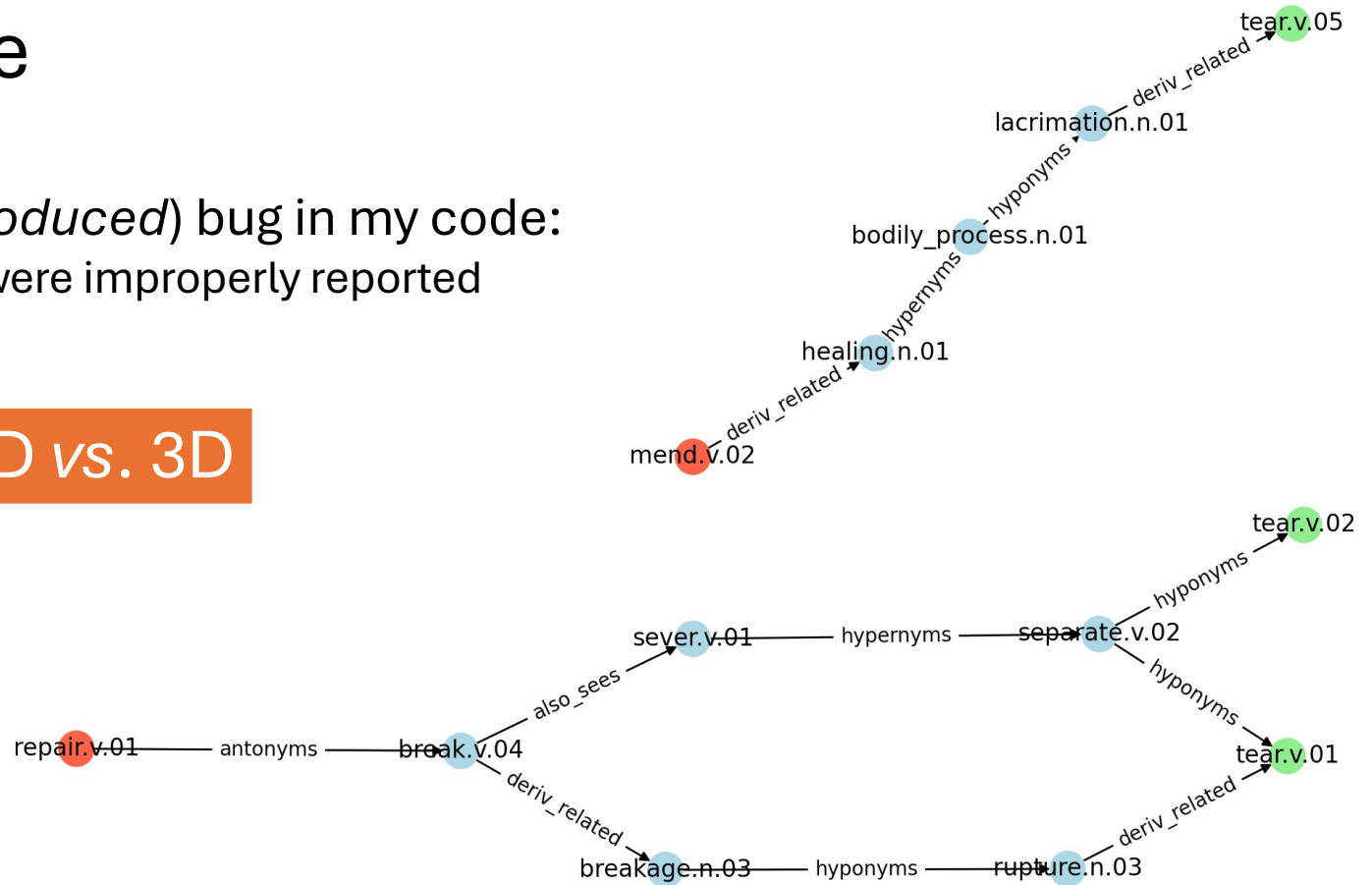
- We have studied:
 - **Word senses:** WordNet, Word Embeddings
 - **nltk:** PTB, also wordnet
 - **Parsers:** dependency and constituency-based
- *Putting things together ...*
- Case study:
 - verb frames of *break*

Last Time

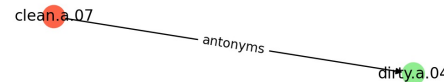
(recently introduced) bug in my code:

- edge labels were improperly reported
- now fixed

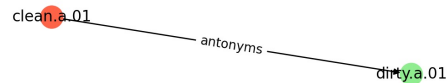
Demo: 2D vs. 3D



clean-dirty



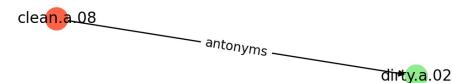
```
>>> wn.synset('clean.a.01').definition()
'free from dirt or impurities; or having clean habits'
>>> wn.synset('clean.a.07').definition()
'not spreading pollution or contamination; especially radioactive
contamination'
>>> wn.synset('clean.a.08').definition()
'(of behavior or especially language) free from objectionable elements;
fit for all observers'
```



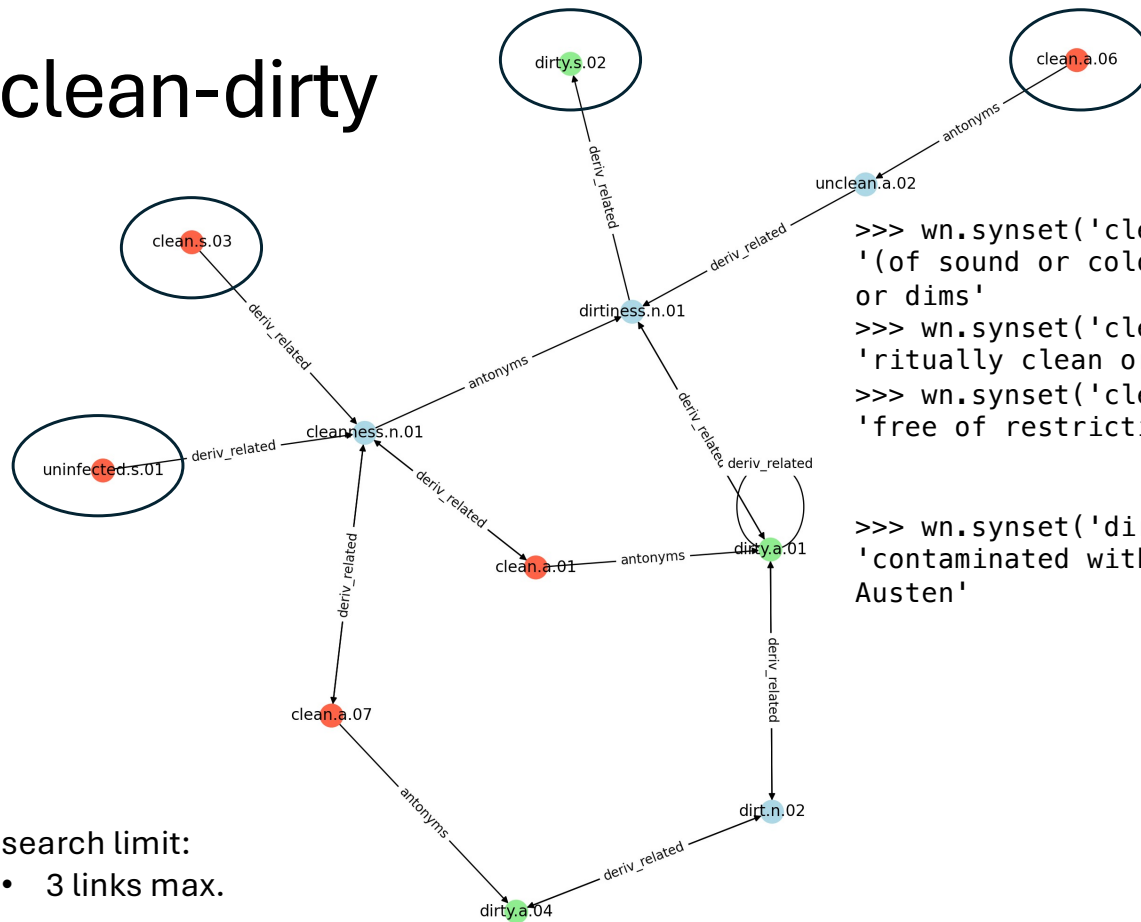
```
>>> wn.synset('dirty.a.01').definition()
'soiled or likely to soil with dirt or grime'
>>> wn.synset('dirty.a.02').definition()
'(of behavior or especially language) characterized by obscenity or indecency'
>>> wn.synset('dirty.a.04').definition()
'spreading pollution or contamination; especially radioactive contamination'
```

search limit:

- 2 links max.



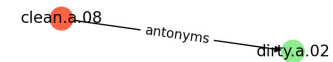
clean-dirty



- search limit:
- 3 links max.

```
>>> wn.synset('clean.a.03').definition()
'(of sound or color) free from anything that dulls
or dims'
>>> wn.synset('clean.a.06').definition()
'ritually clean or pure'
>>> wn.synset('clean.s.01').definition()
'free of restrictions or qualifications'
```

```
>>> wn.synset('dirty.s.02').definition()
'contaminated with infecting organisms; ; - Jane
Austen'
```



Senses of *break*

Lexical Semantics with Large Language Models: A Case Study of English *break**

Erika Petersen
Stanford University
epetsen@stanford.edu

Christopher Potts
Stanford University
cgpotts@stanford.edu

Abstract

Large neural language models (LLMs) can be powerful tools for research in lexical semantics. We illustrate this potential using the English verb *break*, which has numerous senses and appears in a wide range of syntactic frames. We show that LLMs capture known sense distinctions and can be used to identify informative new sense combinations for further analysis. More generally, we argue that LLMs are aligned with lexical semantic theories in providing high-dimensional, contextually modulated representations, but LLMs' lack of discrete features and dependence on usage-based data offer a genuinely new perspective on traditional problems in lexical semantics.

distinctions and can be used to identify new sense combinations for further analysis.

We use these findings as a chance to reflect on the core theoretical commitments of lexical semantics as they pertain to LLM-based investigations. Our discussion is centered around the three tenets of lexical semantics given in Table 1: lexical representations are *high dimensional*, *contextually modulated*, and include *discrete features*.

The high dimensionality property is not phrased as a direct claim in the literature as far as we know, but it reflects the practice of linguists, who identify numerous interacting features of lexical items. Section 2 offers a summary picture for *break*. Similarly, discreteness is often assumed by linguists

Senses of *break*

(Petersen & Potts) <https://ling.auf.net/lingbuzz/006859>

Frame	Sense
1. break the vase	shatter
2. break the computer	render inoperable
3. break the news	reveal
4. break the silence	interrupt
5. break the record	surpass
6. break the code	decipher
7. break the law	violate
8. break the horse	tame
9. break a \$10 bill	make change
10. break the fall	lessen
11. the weather broke	changed
12. the day broke	began

- Change-of-state verb that undergoes the causative alternation:

- *John broke the window*
- *the window broke*

J. hit the tree
**the tree hit*

- "*senses 1–4 all alternate, senses 5–10 are all strictly transitive, ... also strictly intransitive uses, as in 11–12*"

- **the law broke* (present in corpus)
- **careful:** need to expand for context:

"Well," said the man who handed his last cent to the lawyer, "I suppose turnabout is fair play. I broke the law and the law broke me."

Central Law Journal, vol. 36, 1893.

Unaccusative

- Change-of-state verb that undergoes the causative alternation:
 - *John broke the window*
 - *agent causes patient/theme to undergo change of state (become broken)*
 - *the window broke*
 - *window not an agent (usually an external argument)*
 - **the window broke the ball*
 - *window undergoes change of state*
- [_{VP} past [_{VP} *break the window*]] is the underlying structure
 - *the window* raises to subject position (*in the absence of an external argument*) and is pronounced there ("original copy" not pronounced)

More senses of *break*

Frame	Sense
13. break off the engagement	end
14. break out	begin
15. break out in hives	get
16. break into the building	intrude
17. break down the problem	analyze
18. break down the proteins	decompose
19. break in	enter
20. break in	interrupt
21. break free	escape
22. break even	profit = loss
23. break forth	emerge
24. break to the right	turn

- "Except for 13, 17, and 18, these uses are all intransitive"
- Verb particles:
 - *off*
 - *out*
 - *down*
 - *free*
 - *even*
 - *forth*
- particle + NP or PP?:
 - *off the engagement*
 - *out in hives*
 - *into the building*
 - *down the problem/proteins*
 - *to the right*

WordNet: verb senses of *break*

#2-#5

- Verb
- **S: (v) interrupt#4, break#1** (terminate) *"She interrupted her pregnancy"; "break a lucky streak"; "break the cycle of poverty"*
 - **S: (v) break#2, separate#10, split up#4, fall apart#4, come apart#1** (become separated into pieces or fragments) *"The figurine broke"; "The freshly baked loaf fell apart"*
 - **S: (v) break#3** (render inoperable or ineffective) *"You broke the alarm clock when you took it apart!"*
 - **S: (v) break#4, bust#1** (ruin completely) *"He busted my radio!"*
 - **S: (v) break#5** (destroy the integrity of; usually by force; cause to separate into pieces or fragments) *"He broke the glass plate"; "She broke the match"*
 - **S: (v) transgress#1, offend#2, infract#1, violate#2, go against#2, breach#1, break#6** (act in disregard of laws, rules, contracts, or promises) *"offend all laws of humanity"; "violate the basic laws or human civilization"; "break a law"; "break a promise"*
 - **S: (v) break#7, break out#3, break away#1** (move away or escape suddenly) *"The horses broke from the stable"; "Three inmates broke jail"; "Nobody can break out--this prison is high security"*
 - **S: (v) break#8** (scatter or part) *"The clouds broke after the heavy downpour"*
 - **S: (v) break#9, burst#2, erupt#6** (force out or release suddenly and often violently something pent up) *"break into tears"; "erupt in anger"*
 - **S: (v) break#10, break off#2, discontinue#3, stop#6** (prevent completion) *"stop the project"; "break off the negotiations"*
 - **S: (v) break in#1, break#11** (enter someone's (virtual or real) property in an unauthorized manner, usually with the intent to steal or commit a violent act) *"Someone broke in while I was on vacation"; "They broke into my car and stole my radio!"; "who broke into my account last night?"*
 - **S: (v) break in#6, break#12** (make submissive, obedient, or useful) *"The horse was tough to break"; "I broke in the new intern"*
 - **S: (v) violate#1, go against#1, break#13** (fail to agree with; be in violation of; as of rules or patterns) *"This sentence violates the rules of syntax"*
 - **S: (v) better#1, break#14** (surpass in excellence) *"She bettered her own record"; "break a record"*

- **S: (v) unwrap#2, disclose#1, let on#1, bring out#9, reveal#2, discover#6, expose#2, divulge#1, break#15, give away#2, let out#2, uncover#3** (make known to the public information that was previously known only to a few people or that was meant to be kept a secret) *"The auction house would not disclose the price at which the van Gogh had sold"; "The actress won't reveal how old she is"; "bring out the truth"; "he broke the news to her"; "unwrap the evidence in the murder case"; "The newspaper uncovered the President's illegal dealings"*
- **S: (v) break#16** (come into being) *"light broke over the horizon"; "Voices broke in the air"*
- **S: (v) fail#4, go bad#1, give way#4, die#4, give out#4, conk out#1, go#30, break#17, break down#4** (stop operating or functioning) *"The engine finally went"; "The car died on the road"; "The bus we travelled in broke down on the way to town"; "The coffee maker broke"; "The engine failed on the way to town"; "her eyesight went after the accident"*
- **S: (v) break#18, break away#3** (interrupt a continued activity) *"She had broken with the traditional patterns"*
- **S: (v) break#19** (make a rupture in the ranks of the enemy or one's own by quitting or fleeing) *"The ranks broke"*
- **S: (v) break#20** (curl over and fall apart in surf or foam, of waves) *"The surf broke"*
- **S: (v) dampen#7, damp#4, soften#2, weaken#5, break#21** (lessen in force or effect) *"soften a shock"; "break a fall"*
- **S: (v) break#22** (be broken in) *"If the new teacher won't break, we'll add some stress"*
- **S: (v) break#23** (come to an end) *"The heat wave finally broke yesterday"*
- **S: (v) break#24** (vary or interrupt a uniformity or continuity) *"The flat plain was broken by tall mesas"*
- **S: (v) break#25** (cause to give up a habit) *"She finally broke herself of smoking cigarettes"*
- **S: (v) break#26** (give up) *"break cigarette smoking"*
- **S: (v) break#27** (come forth or begin from a state of latency) *"The first winter storm broke over New York"*
- **S: (v) break#28** (happen or take place) *"Things have been breaking pretty well for us in the past few months"*
- **S: (v) break#29** (cause the failure or ruin of) *"His peccadilloes finally broke his marriage"; "This play will either make or break the playwright"*
- **S: (v) break#30** (invalidate by judicial action) *"The will was broken"*

WordNet: verb senses of *break*

- [S](#): (v) [separate#8](#), [part#2](#), [split up#3](#), [split#3](#), [break#31](#), [break up#2](#) (discontinue an association or relation; go different ways) "The business partners broke over a tax question"; "The couple separated after 25 years of marriage"; "My friend and I split up"
- [S](#): (v) [demote#1](#), [bump#4](#), [relegate#2](#), [break#32](#), [kick downstairs#1](#) (assign to a lower position; reduce in rank) "She was demoted because she always speaks up"; "He was broken down to Sergeant"
- [S](#): (v) [bankrupt#1](#), [ruin#3](#), [break#33](#), [smash#3](#) (reduce to bankruptcy) "My daughter's fancy wedding is going to break me!"; "The slump in the financial markets smashed him"
- [S](#): (v) [break#34](#) (change directions suddenly)
- [S](#): (v) [break#35](#) (emerge from the surface of a body of water) "The whales broke"
- [S](#): (v) [collapse#1](#), [fall in#1](#), [cave in#1](#), [give#26](#), [give way#2](#), [break#36](#), [founder#3](#) (break down, literally or metaphorically) "The wall collapsed"; "The business collapsed"; "The dam broke"; "The roof collapsed"; "The wall gave in"; "The roof finally gave under the weight of the ice"
- [S](#): (v) [break dance#1](#), [break-dance#1](#), [break#37](#) (do a break dance) "Kids were break-dancing at the street corner"
- [S](#): (v) [break#38](#) (exchange for smaller units of money) "I had to break a \$100 bill just to buy the candy"
- [S](#): (v) [break#39](#), [break up#9](#) (destroy the completeness of a set of related items) "The book dealer would not break the set"
- [S](#): (v) [break#40](#) (make the opening shot that scatters the balls)
- [S](#): (v) [break#41](#) (separate from a clinch, in boxing) "The referee broke the boxers"
- [S](#): (v) [break#42](#), [wear#7](#), [wear out#2](#), [bust#4](#), [fall apart#2](#) (go to pieces) "The lawn mower finally broke"; "The gears wore out"; "The old chair finally fell apart completely"
- [S](#): (v) [break#43](#), [break off#4](#), [snap off#1](#) (break a piece from a whole) "break a branch from a tree"
- [S](#): (v) [break#44](#) (become punctured or penetrated) "The skin broke"
- [S](#): (v) [break#45](#) (pierce or penetrate) "The blade broke her skin"
- [S](#): (v) [break#46](#), [get out#6](#), [get around#1](#) (be released or become known; of news) "News of her death broke in the morning"
- [S](#): (v) [pause#2](#), [intermit#1](#), [break#47](#) (cease an action temporarily) "We pause for station identification"; "let's break for lunch"
- [S](#): (v) [break#48](#) (interrupt the flow of current in) "break a circuit"

- [S](#): (v) [break#49](#) (undergo breaking) "The simple vowels broke in many Germanic languages"
- [S](#): (v) [break#50](#) (find a flaw in) "break an alibi"; "break down a proof"
- [S](#): (v) [break#51](#) (find the solution or key to) "break the code"
- [S](#): (v) [break#52](#) (change suddenly from one tone quality or register to another) "Her voice broke to a whisper when she started to talk about her children"
- [S](#): (v) [break#53](#), [recrudescence#1](#), [develop#20](#) (happen) "Report the news as it develops"; "These political movements recrudescence from time to time"
- [S](#): (v) [crack#1](#), [check#25](#), [break#54](#) (become fractured; break or crack on the surface only) "The glass cracked when it was heated"
- [S](#): (v) [break#55](#) (crack; of the male voice in puberty) "his voice is breaking--he should no longer sing in the choir"
- [S](#): (v) [break#56](#) (fall sharply) "stock prices broke"
- [S](#): (v) [fracture#6](#), [break#57](#) (fracture a bone of) "I broke my foot while playing hockey"
- [S](#): (v) [break#58](#) (diminish or discontinue abruptly) "The patient's fever broke last night"
- [S](#): (v) [break#59](#) (weaken or destroy in spirit or body) "His resistance was broken"; "a man broken by the terrible experience of near-death"

down a proof

to a whisper

unaccusative:
subject is affected

for lunch

WordNet: senses of *break out* and *break down*

Verb

- **S: (v) erupt#1, break out#1** (start abruptly) *"After 1989, peace broke out in the former East Bloc"*
- **S: (v) break out#2** (begin suddenly and sometimes violently) *"He broke out shouting"*
- **S: (v) break#7, break out#3, break away#1** (move away or escape suddenly) *"The horses broke from the stable"; "Three inmates broke jail"; "Nobody can break out--this prison is high security"*
- **S: (v) break out#4** (take from stowage in preparation for use)
- **S: (v) erupt#8, recrudescence#2, break out#5** (become raw or open) *"He broke out in hives"; "My skin breaks out when I eat strawberries"; "Such boils tend to recrudescence"*

Verb

- **S: (v) break down#1, crush#7** (make ineffective) *"Martin Luther King tried to break down racial discrimination"*
- **S: (v) analyze#2, analyse#3, break down#2, dissect#2, take apart#3** (make a mathematical, chemical, or grammatical analysis of; break down into components or essential features) *"analyze a specimen"; "analyze a sentence"; "analyze a chemical compound"*
- **S: (v) break down#3, lose it#1, snap#11** (lose control of one's emotions) *"When she heard that she had not passed the exam, she lost it completely"; "When her baby died, she snapped"*
- **S: (v) fail#4, go bad#1, give way#4, die#4, give out#4, conk out#1, go#30, break#17, break down#4** (stop operating or functioning) *"The engine finally went"; "The car died on the road"; "The bus we travelled in broke down on the way to town"; "The coffee maker broke"; "The engine failed on the way to town"; "her eyesight went after the accident"*
- **S: (v) crumble#1, crumple#1, tumble#5, break down#5, collapse#4** (fall apart) *"the building crumbled after the explosion"; "Negotiations broke down"*
- **S: (v) break down#6** (cause to fall or collapse)
- **S: (v) decompose#1, break up#18, break down#7** (separate (substances) into constituent elements or parts)
- **S: (v) break down#8, collapse#2** (collapse due to fatigue, an illness, or a sudden attack)

WordNet: senses of *break in* and *break off*

Verb

- [S: \(v\) break in#1](#), [break#11](#) (enter someone's (virtual or real) property in an unauthorized manner, usually with the intent to steal or commit a violent act) "*Someone broke in while I was on vacation*"; "*They broke into my car and stole my radio!*"; "*who broke into my account last night?*"
- [S: \(v\) chime in#1](#), [cut in#3](#), [put in#3](#), [butt in#1](#), [chisel in#1](#), [barge in#2](#), [break in#2](#) (break into a conversation) "*her husband always chimes in, even when he is not involved in the conversation*"
- [S: \(v\) break in#3](#) (start in a certain activity, enterprise, or role)
- [S: \(v\) break in#4](#) (intrude on uninvited) "*The nosy couple broke in on our conversation*"
- [S: \(v\) break in#5](#) (break so as to fall inward) "*He broke in the door*"
- [S: \(v\) break in#6](#), [break#12](#) (make submissive, obedient, or useful) "*The horse was tough to break*"; "*I broke in the new intern*"

Verb

- [S: \(v\) cut short#1](#), [break short#1](#), [break off#1](#) (interrupt before its natural or planned end) "*We had to cut short our vacation*"
- [S: \(v\) break#10](#), [break off#2](#), [discontinue#3](#), [stop#6](#) (prevent completion) "*stop the project*"; "*break off the negotiations*"
- [S: \(v\) chip#1](#), [chip off#1](#), [come off#3](#), [break away#2](#), [break off#3](#) (break off (a piece from a whole)) "*Her tooth chipped*"
- [S: \(v\) break#43](#), [break off#4](#), [snap off#1](#) (break a piece from a whole) "*break a branch from a tree*"
- [S: \(v\) chip#5](#), [knap#2](#), [cut off#5](#), [break off#5](#) (break a small piece off from) "*chip the glass*"; "*chip a tooth*"

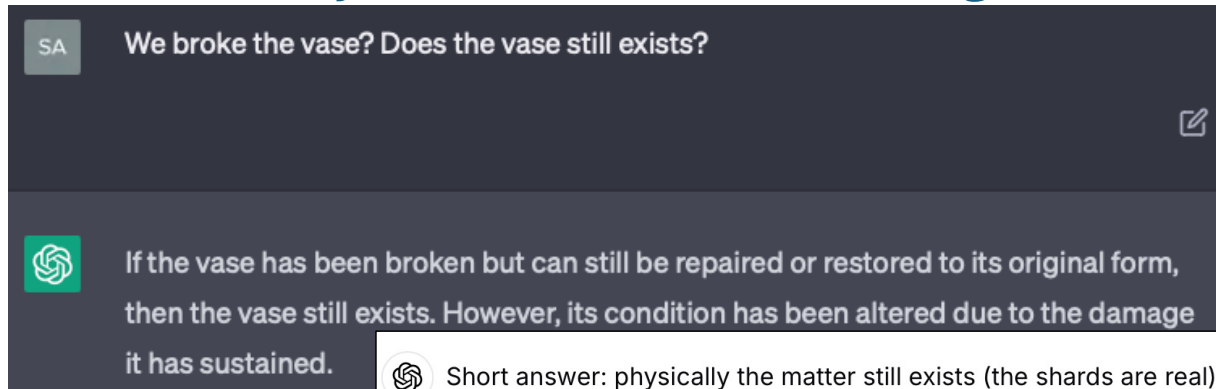
Let's ask ChatGPT (3.5 & 5 Mini)

- Verb synsets are arranged into hierarchies as well
 - depends on the **semantic field**: *volume, speed, intensity of emotion etc.*
- break > pulverize > obliterate (*my intuition*)


```
>>> wn.synsets('pulverize')
[Synset('powderize.v.02'), Synset('demolish.v.01'), Synset('powderize.v.01')]
>>> wn.synsets('obliterate')
[Synset('kill.v.12'), Synset('obscure.v.05'), Synset('obliterate.v.03'),
Synset('obliterate.v.04'), Synset('blotted_out.s.01')]
```

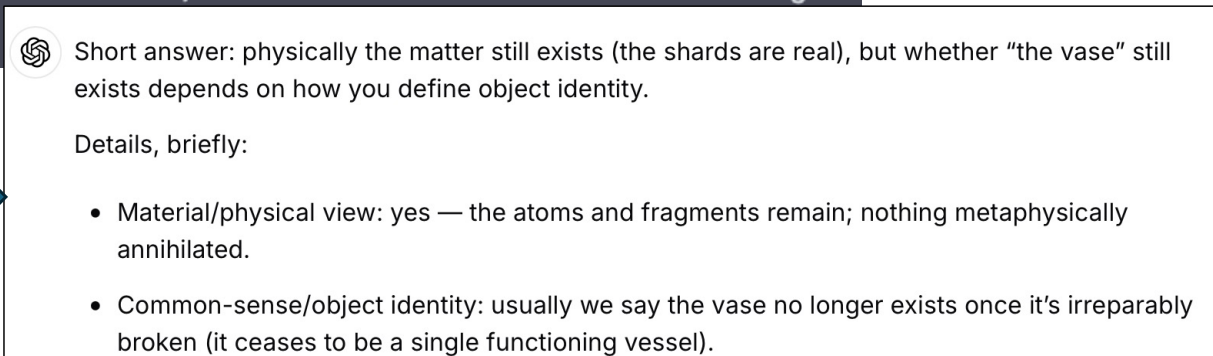

Let's ask ChatGPT (3.5 & 5 Mini)


- *Does the object still exist after being broken?*



SA We broke the vase? Does the vase still exists?

 If the vase has been broken but can still be repaired or restored to its original form, then the vase still exists. However, its condition has been altered due to the damage it has sustained.



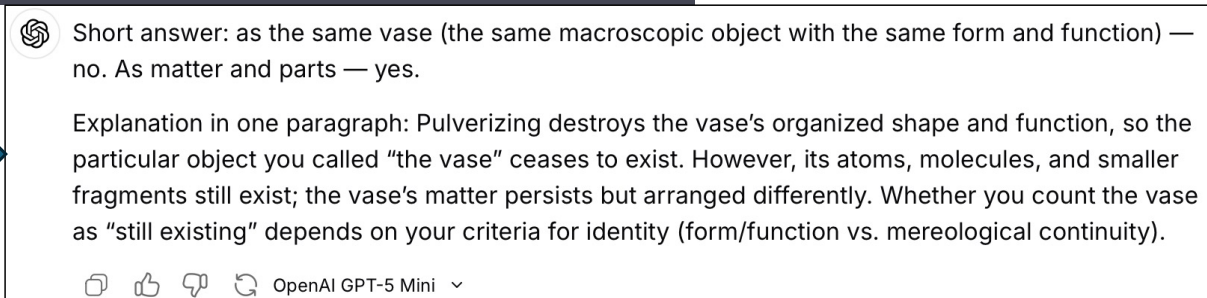
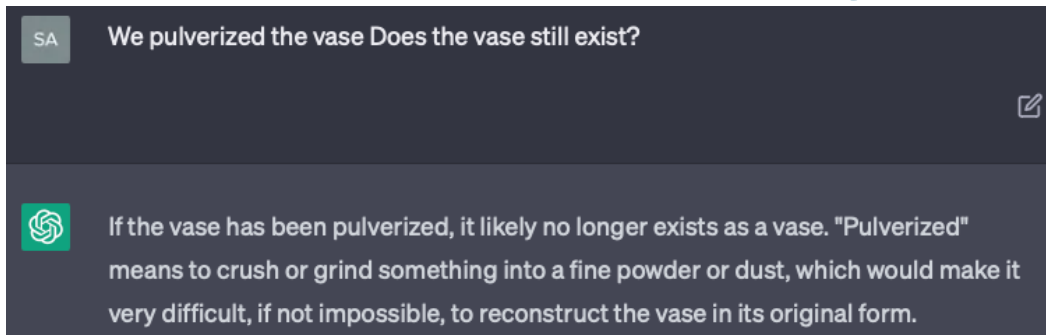
 Short answer: physically the matter still exists (the shards are real), but whether “the vase” still exists depends on how you define object identity.

Details, briefly:

- Material/physical view: yes — the atoms and fragments remain; nothing metaphysically annihilated.
- Common-sense/object identity: usually we say the vase no longer exists once it’s irreparably broken (it ceases to be a single functioning vessel).

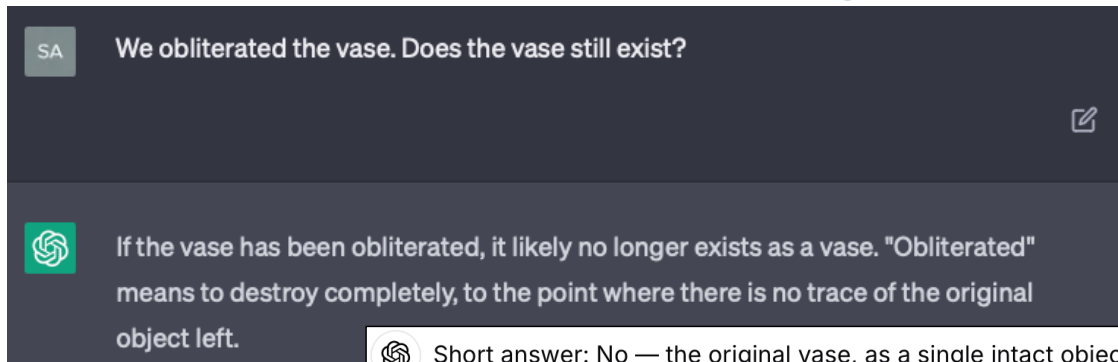
Let's ask ChatGPT (3.5 & 5 Mini)

- *Does the object still exist after being broken?*



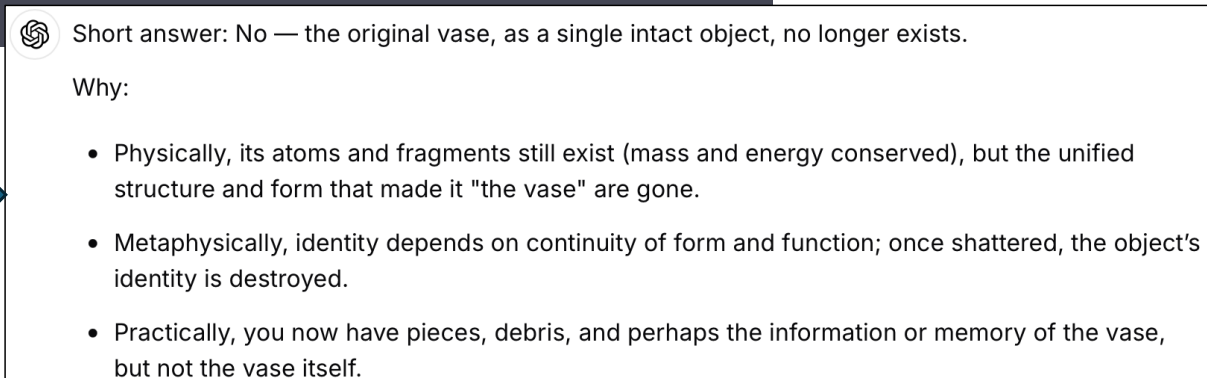
Let's ask ChatGPT

- *Does the object still exist after being broken?*



SA We obliterated the vase. Does the vase still exist?

If the vase has been obliterated, it likely no longer exists as a vase. "Obliterated" means to destroy completely, to the point where there is no trace of the original object left.



Short answer: No — the original vase, as a single intact object, no longer exists.

Why:

- Physically, its atoms and fragments still exist (mass and energy conserved), but the unified structure and form that made it "the vase" are gone.
- Metaphysically, identity depends on continuity of form and function; once shattered, the object's identity is destroyed.
- Practically, you now have pieces, debris, and perhaps the information or memory of the vase, but not the vase itself.

nltk ptb

- Recall we installed this a while ago:

```
>>> from nltk.corpus import ptb
>>> type(ptb)
<class 'nltk.corpus.util.LazyCorpusLoader'>
>>> ptb.fileids()
['BROWN/CF/CF01.MRG', 'BROWN/CF/CF02.MRG', ... , 'WSJ/24/WSJ_2453.MRG', 'WSJ/24/WSJ_2454.MRG']
>>> len(ptb.fileids())
2504
>>> len(ptb.parsed_sents())
73451
>>> len(ptb.words())
1740895
```

nltk ptb.parsed_sents()

- nltk Tree objects:

- <https://www.nltk.org/api/nltk.tree.tree.html>

- Tree printing:

- `t.pretty_print()` ascii graphics
- `t.pprint()` parentheses
- `t.draw()` window

Note: Stanza Tree objects are different

- *more on this next time ...*

- Recall Tree objects are recursively defined. Given a nltk tree t:

- `t.label()` category label of t
- `len(t)` number of children of t
- `t[i]` ith child of t, could be Tree or string: `isinstance(t[i], str)`
- `t.subtrees()` all possible subtrees in t
- `t.pos()` list of (*word*, *postag*) in t

← *be using this one first*

Find all ptb *break* sentences

break.py

```
1 from nltk.corpus import ptb
2
3 verbforms = set(['break', 'breaks', 'broke', 'breaking', 'broken'])
4
5 def has_break(list):
6     for wordtag in list:
7         if wordtag[0] in verbforms and wordtag[1].startswith('VB'):
8             return True
9     return False
10
11 break_trees = [t for t in ptb.parsed_sents() if has_break(t.pos())]
```

Find all ptb *break* sentences

```
$ python3 -i break.py  
>>> len(break_trees)  
244
```



run code, drop into interpreter

- Let's write Python code to show the phrases in which *break* appears!

Programming: all ptb *break* phrases

- **Idea:**

- use nltk Tree methods to drill down until
- find phrase that immediately dominates a node with pos tag VB*
- node must immediately dominate (a form of) *break* in verbforms

- **Code sketch:**

- loop over all break_trees (from break.py)
 - loop over each subtree st `t.subtrees()` all possible subtrees in t
 - if st not a string
 - loop over each child stc of st
 - if stc not a string
 - if stc has pos tag VB* and 0th child of stc in verbforms
 - print st

`t[i]` *i*th child of t, Tree or string
`len(t)` number of children of t
`isinstance(t[i],str)`

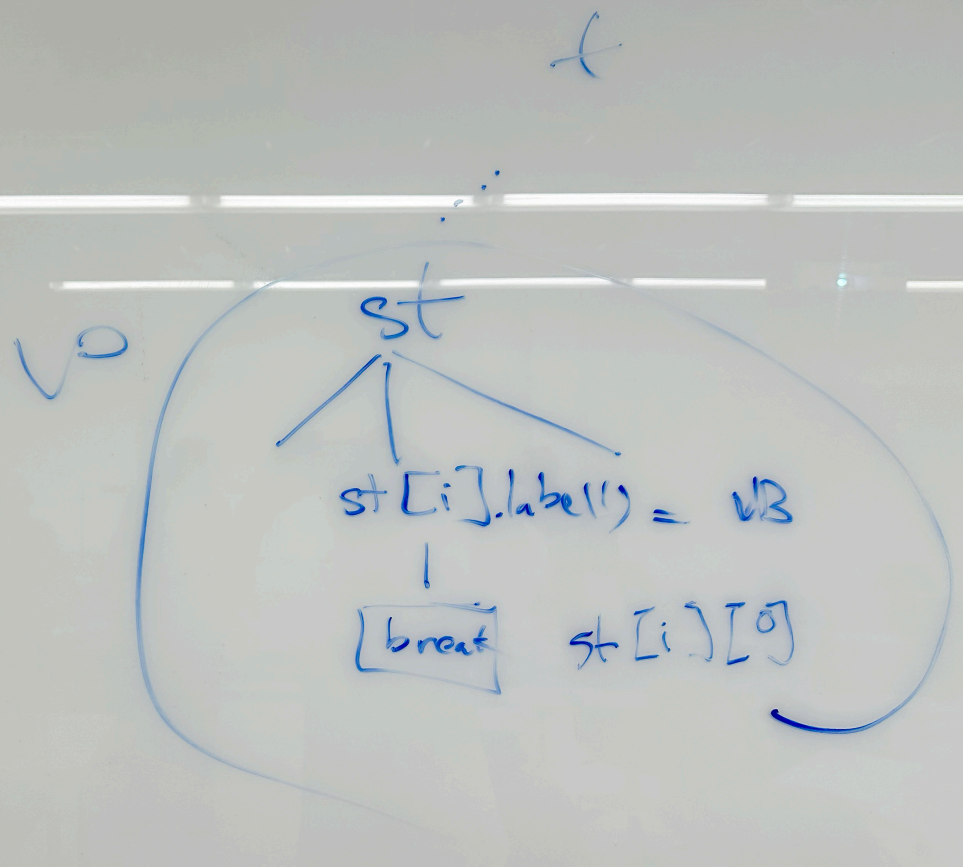
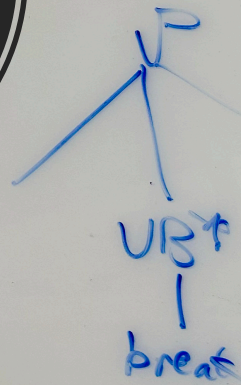
Programming: all ptb *break* phrases

Code output

- *note: instead of printing, we could accumulate these subtrees into a list*

```
(NP (DT a) (JJ certain) (VBN broken) (NN relic))
(NP (VBN broken) (NNS bridges))
(NP (DT the) (VBN broken) (NN railroad))
(VP (VB break) (ADVP (IN through)) (SBAR-LOC (WHADVP-2 (WRB where)) (S (NP-SBJ (NP (DT some) (NN feature)) (PP (IN of) (NP (DT the)
(NN winter) (NN landscape)))) (VP (MD would) (VP (VB be) (ADJP-PRD (JJ recognizable) (ADVP-LOC (-NONE- *1*-2)))))))
(VP (VBD broke) (NP (DT the) (NN language) (NN barrier)))
(VP (VBN broken) (PP (IN with) (NP (PRP$ his) (NNS parents))) (SBAR-PRP (IN in) (NN order) (S (NP-SBJ (-NONE- *-3)) (VP (TO to) (VP
(VB keep) (NP (NP (NP (DT the) (NN company)) (SBAR (WHNP-2 (-NONE- 0)) (S (NP-SBJ (PRP he)) (VP (VBD preferred) (NP (-NONE- *T*-
2)))))) (: --) (NP (NNS pimps) (, ,) (NNS panders) (CC and) (NNS whores)))))))
(VP (VBD broke) (PRT (RP down)) (PP (IN into) (NP (DT a) (JJ general) (, ,) (JJ wordy) (NN argument))))
(VP (VBD broke) (PRT (RP up)))
(VP (VBG breaking) (NP (DT the) (JJ thick) (NN sod)))
(VP (VB break) (PRT (RP up)) (NP (NP (DT the) (NN whole)) (PP (IN of) (NP (DT that) (JJ unfortunate) (NN establishment)))))
(VP (VB break) (NP (-NONE- *T*-2)))
(VP (VBP break))
(VP (VBG breaking) (, ,) (SBAR-ADV (IN as) (S (NP-SBJ (JJ rigid) (NN surface) (NNS rocks)) (VP (VBP do))))
(VP (VBD broke) (PRT (RP out)))
(VP (VBD broke) (PP (PP (IN into) (NP (DT a) (NN trot))) (CONJP (CC and) (RB then)) (PP (IN into) (NP (DT a) (NN run)))))
(VP (VBD broke) (PRT (RP up)) (NP (NP (DT the) (NN timing)) (PP (IN of) (NP (DT the) (JJ entire) (NN enemy) (NN attack)))))
(VP (VBN broken) (NP (-NONE- *-3)) (PP (IN for) (NP (NP (DT the) (NN superintendent) (POS 's)) (NN home))))
```

Whiteboard

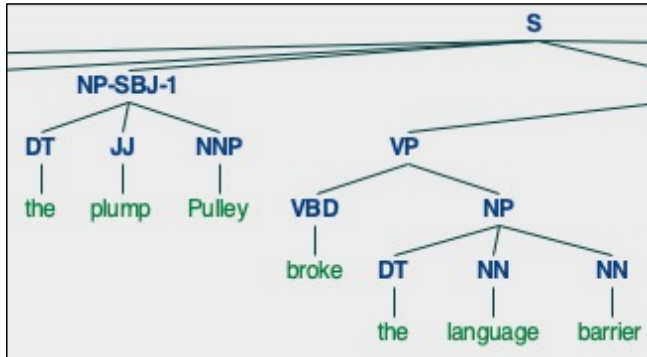
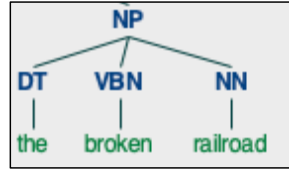
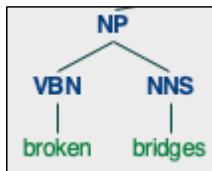
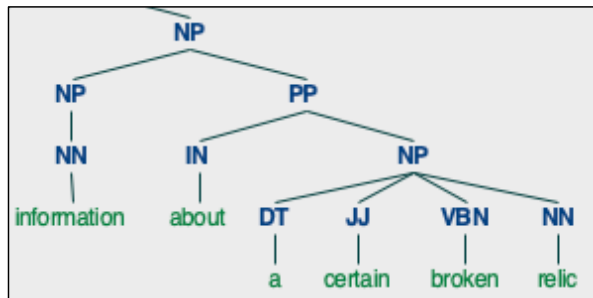


Find all *break* sentences

- Could draw all trees of course:

```
>>> break_trees[0].draw()
```

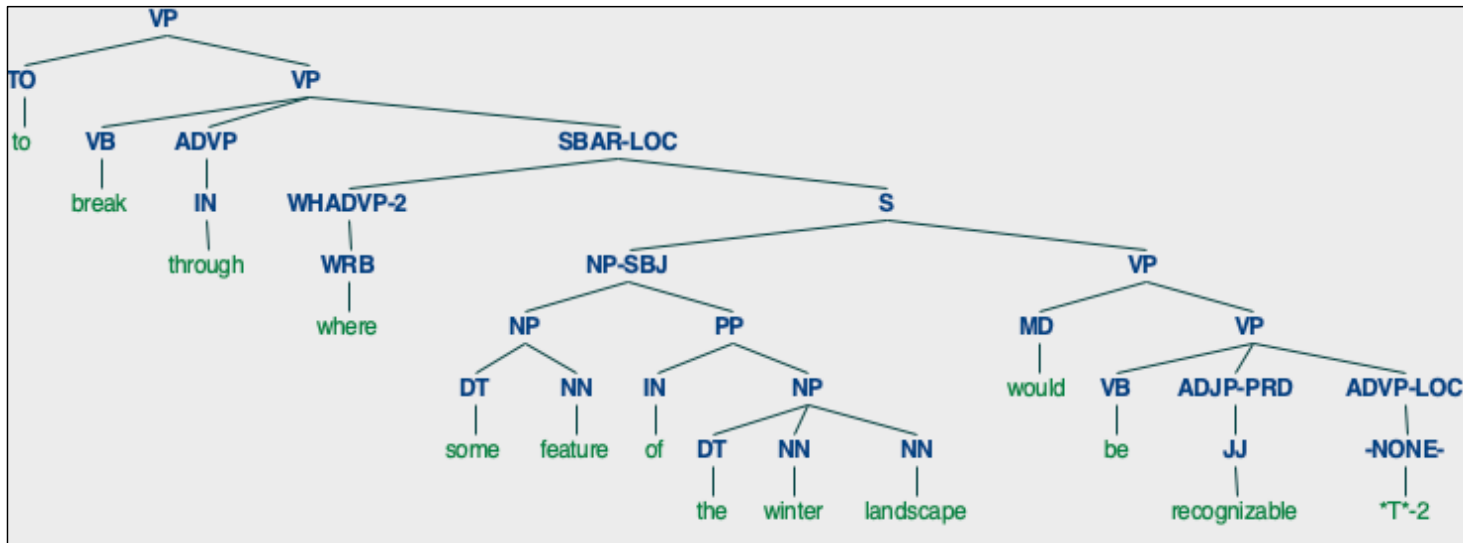
adjectival modification with VBN *broken*



break NP

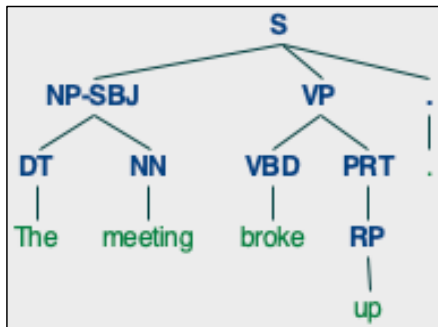
Find all *break* sentences

break through/ADVP



Find all *break* sentences

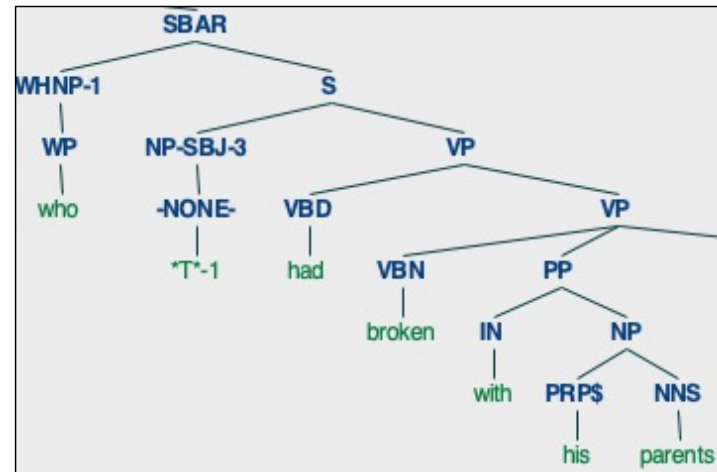
break up/PRT



RP

Particle

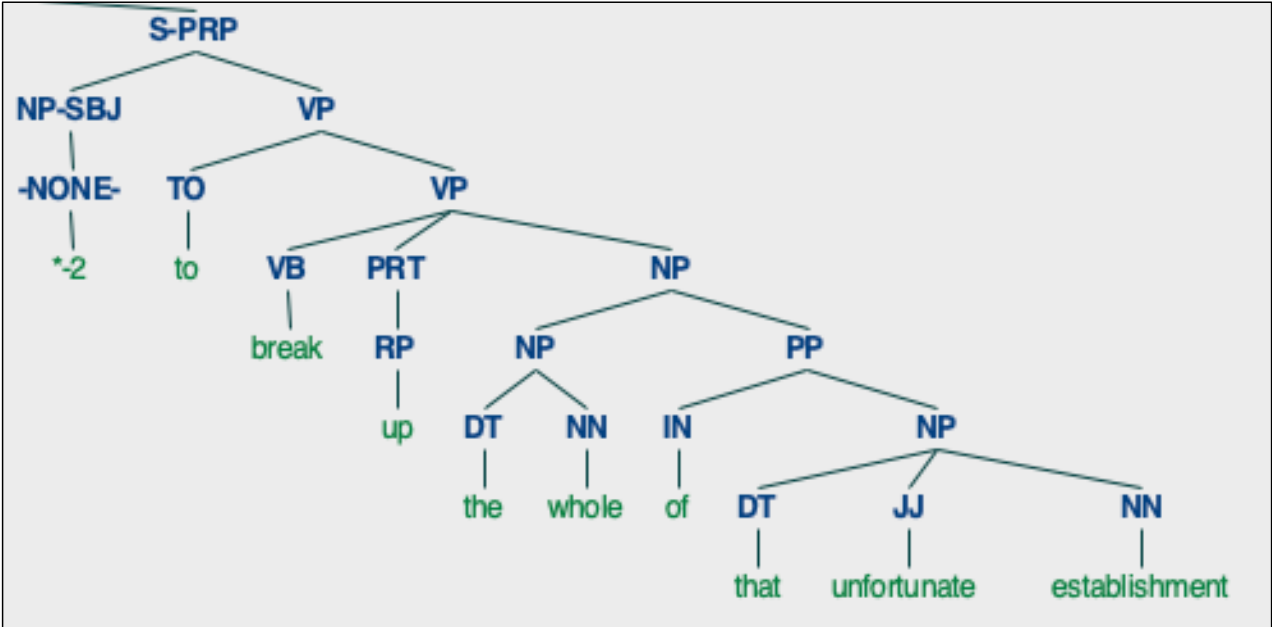
break PP[with]



Find all *break* sentences

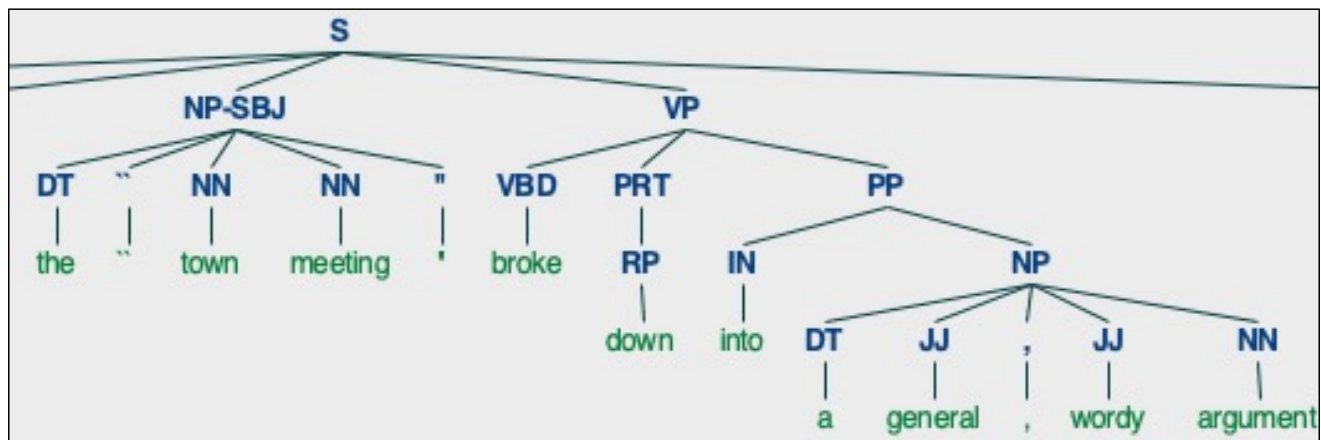
break up/PRT NP

RP Particle



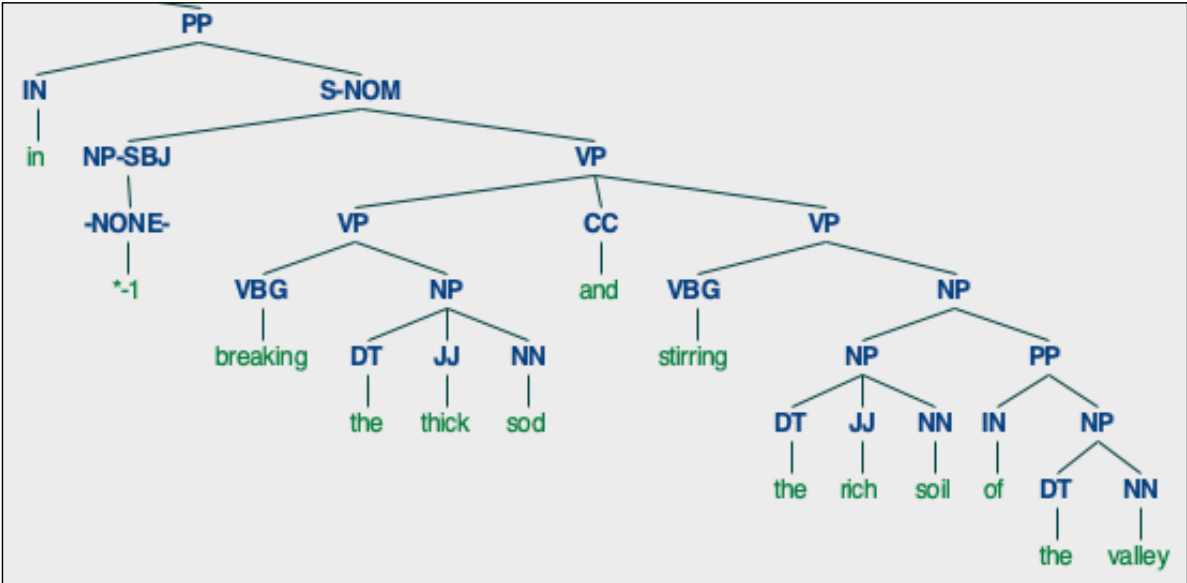
Find all *break* sentences

break down/PRT PP[*into*]



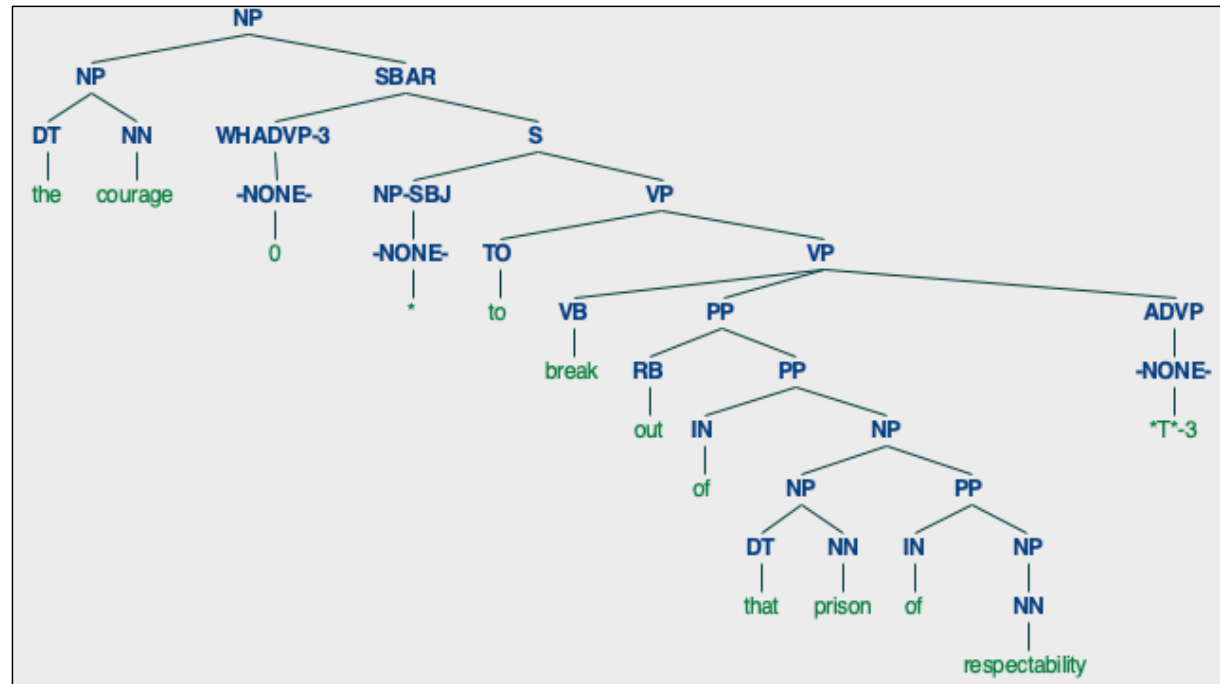
RP Particle

Find all *break* sentences



Find all *break* sentences

break out of NP



Programming: all ptb *break* phrases

- We are interested in verb frames only (VP): *how many are there?*

(NP (DT a) (JJ certain) (VBN broken) (NN relic))

(NP (VBN broken) (NNS bridges))

(NP (DT the) (VBN broken) (NN railroad))

(VP (VB break) (ADVP (IN through)) (SBAR-LOC ...))

(VP (VBD broke) (NP (DT the) (NN language) (NN barrier)))

(VP (VBN broken) (PP (IN with) (NP (PRP\$ his) (NNS parents)))(SBAR-PRP ...))

(VP (VBD broke) (PRT (RP down)) (PP (IN into) (NP ...)))

(VP (VBD broke) (PRT (RP up)))

(VP (VBG breaking) (NP (DT the) (JJ thick) (NN sod)))

(VP (VB break) (PRT (RP up)) (NP (NP (DT the) (NN whole)) (PP ...))))

(VP (VB break) (NP (-NONE- *T*-2)))

(VP (VBP break))

...

Programming: all ptb *break* phrases

- Let's write Python code to keep only the VP frames!
 - `subtree.label().startswith('VP')`
 - also, instead of printing, let save them into a list of Trees `break_vfs`

```
>>> len(break_vfs)
224
```

```
$ python3 -i break.py
>>> len(break_trees)
```

```
244
```

```
(VP (VB break) (ADVP (IN through) (NP (DT the) (NN feature))) (PP (IN of) (NP (DT the) (NN PRD (JJ recognizable)) (ADVP-LUC (-NONE- *I*-2)))))) (S (NP-SBJ (NP (DT some) (NN feature))) (VP (MD would) (VP (VB be) (ADJP-PRD (JJ recognizable)) (ADVP-LUC (-NONE- *I*-2))))))
(VP (VBD broke) (NP (DT the) (NN language) (NN barrier)))
(VP (VBN broken) (PP (IN with) (NP (PRP$ his) (NNS parents))) (SBAR-PRP (IN in) (NN order) (S (NP-SBJ (-NONE- *-3)) (VP (TO to) (VP (VB keep) (NP (NP (NP (DT the) (NN company)) (SBAR (WHNP-2 (-NONE- 0)) (S (NP-SBJ (PRP he)) (VP (VBD preferred) (NP (-NONE- *T*-2)))))) (: --) (NP (NNS pimps) (, ,) (NNS panders) (CC and) (NNS whores))))))))))
(VP (VBD broke) (PRT (RP down)) (PP (IN into) (NP (DT a) (JJ general) (, ,) (JJ wordy) (NN argument))))
(VP (VBD broke) (PRT (RP up)))
(VP (VBG breaking) (NP (DT the) (JJ thick) (NN sod)))
...
```

