

Integrating Intransitive Phrasal Verbs into a NooJ Dictionary

Peter A. Machonis

Florida International University

multiword expressions

- A key problem that must be resolved for NLP to succeed
- English phrasal verbs are some of the most complex cases
- Particles can be easily mistaken with prepositions
- Many phrasal verbs also involve insertions

Some simple examples

Phrasal verb

- We **looked up** the information
- We **looked** the information **up**
- Ann **turned on** the computer
- Ann **turned** the computer **on**

Verb + prepositional phrase

- Max looked **up Mary's dress**
- *Max **looked** Mary's dress **up**
- The bus driver turned **on 48th St**
- *The bus driver **turned** 48th St. **on**

Some real discontinuous examples

- *She had **reasoned** the matter well **out***
- ... the Clinton administration that really **turned** the national economy **around?**
- ... a program that has effectively **brought** our crime rates **down**, our violent crime rates **down**.

Present NooJ dictionary of Phrasal verbs

NooJ

File Edit Lab Project Windows Info DICTIONARY

Dictionary contains 1256 entries

Entry	S-Lemma	Category	FLX	FR	Nlexample
beef	beef	V	ASK	renforcer	the proposal
bend	bend	V	BEND	tordre	the credit card
bind	bind	V	FIND	bander	the wound
block	block	V	ASK	obstruer	the sink
blow	blow	V	BLOW	agrandir	the photo
blow	blow	V	BLOW	exagérer	the scandal
blow	blow	V	BLOW	exploser	the building
blow	blow	V	BLOW	gonfler	the balloons
boil	boil	V	ASK	bouillir	some water

14 sec Cancel

NooJ

File Edit Lab Project Windows Info DICTIONARY

Dictionary contains 1256 entries

Part	synonym	Syntax	SynSem
up	strengthen	-	PV+FXC+NOHum+NONHum+N1NHum
up	bend completely	-	PV+FXC+NOHum+NONHum+N1NHum+NoPartPart
up	put bandage on	-	PV+FXC+NOHum+N1Hum+N1NHum+NoPartPart
up	obstruct	-	PV+FXC+NOHum+NONHum+N1NHum+NoPart+N1VPart
up	enlarge	-	PV+FXC+NOHum+NONHum+N1NHum
up	exaggerate	-	PV+FXC+NOHum+NONHum+N1NHum+N1VPart
up	explode	-	PV+FXC+NOHum+NONHum+N1Hum+N1NHum+N1VPart+N1V
up	inflate	-	PV+FXC+NOHum+NONHum+N1NHum
up	boil	-	PV+FXC+NOHum+N1NHum+NoPartPart+N1V

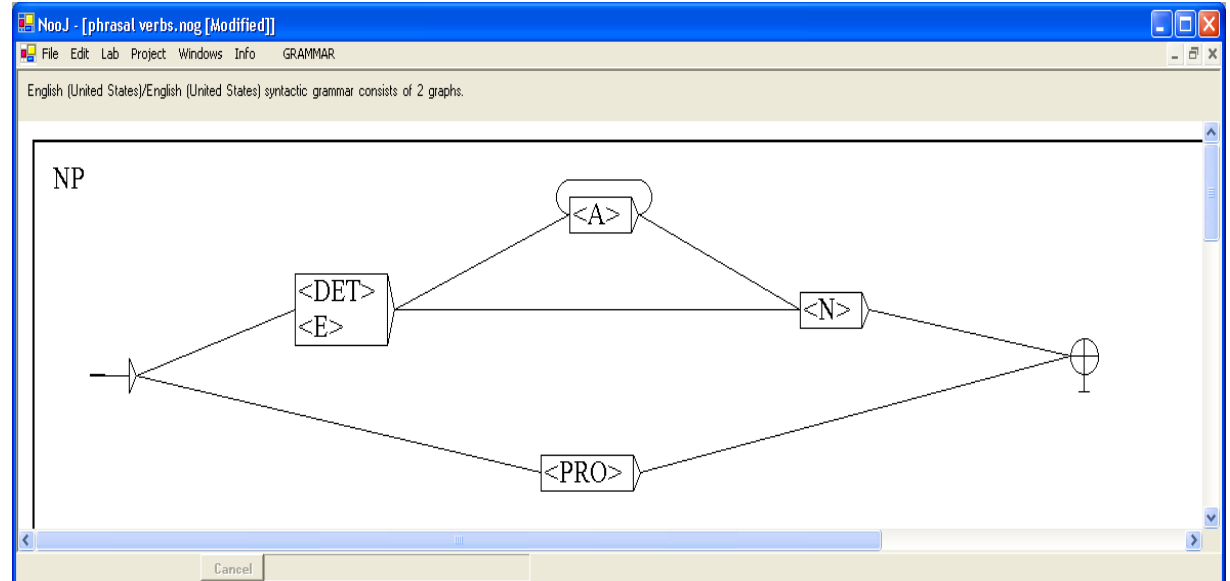
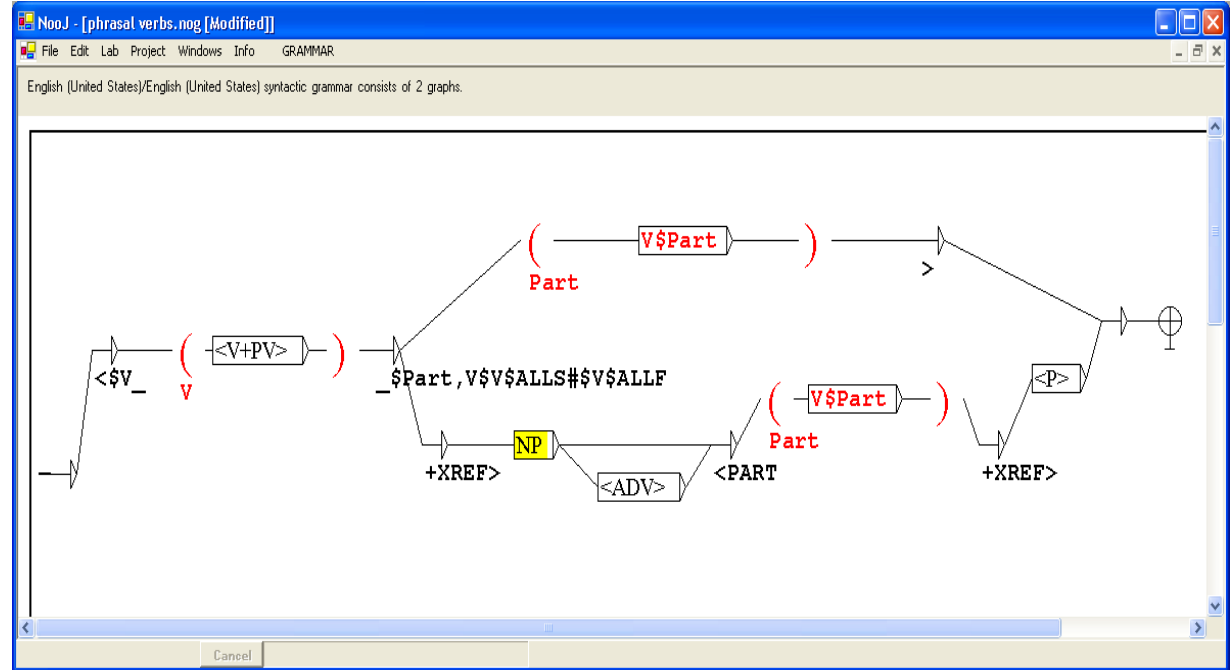
14 sec Cancel

Based on Lexicon-Grammar tables

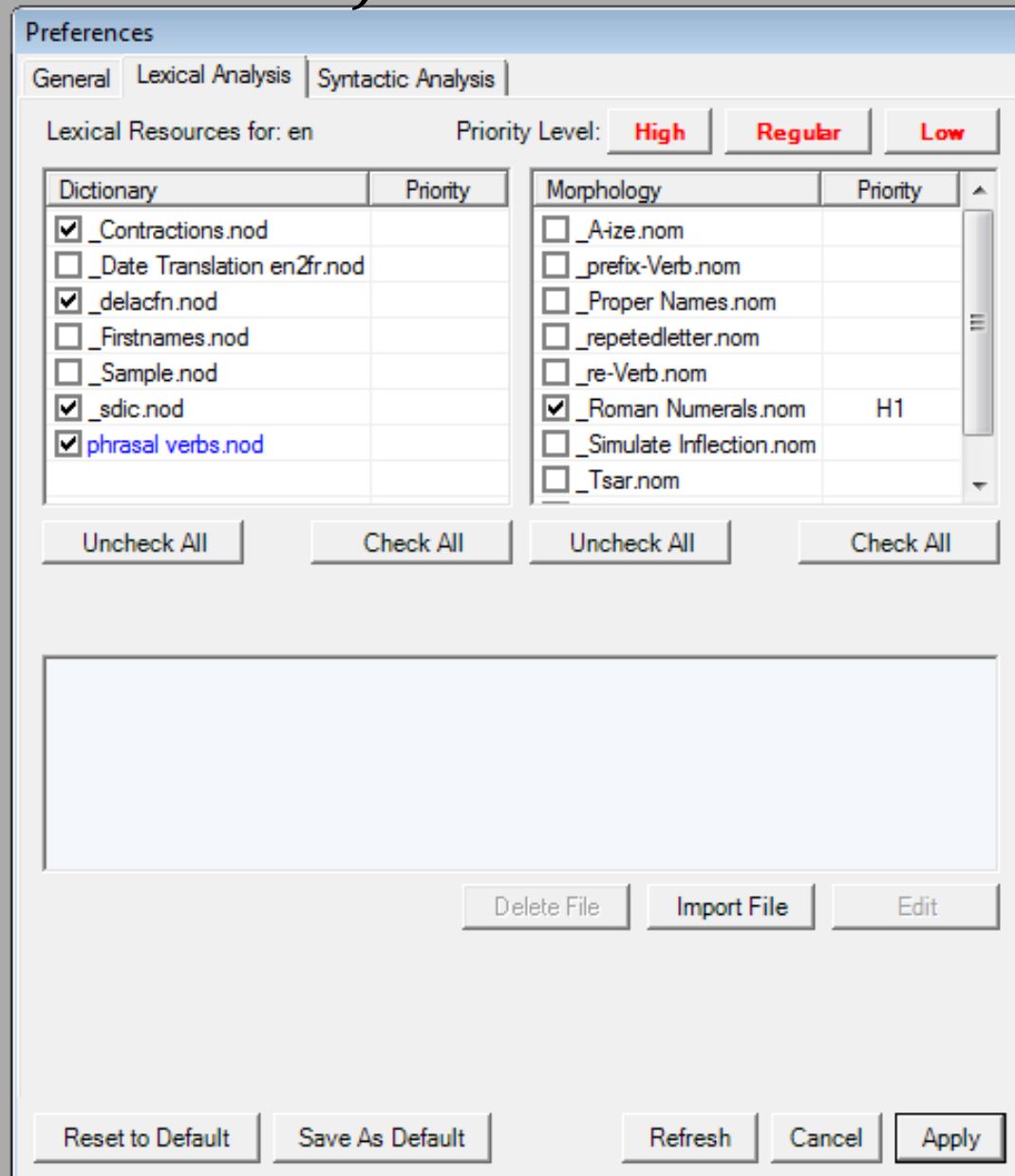
$N_0 ::= N_{hum}$	$N_0 ::= N\text{-}hum$	Verb	Particle	Example of N_1	$N_1 ::= N_{hum}$	$N_1 ::= N\text{-}hum$	$N_0 V N_1$	$N_1 V$ Part	$N_1 V$	Synonym
+	+	beam	up	the aliens	+	+	-	+	-	transport by energy
+	+	bear	up	the weight	+	+	+	-	-	support
+	+	beat	up	the door	-	+	-	-	-	damage
+	+	beat	up	the eggs	-	+	+	-	-	beat
+	-	beat	up	the child	+	-	+	-	-	attack physically & hurt
+	+	beef	up	the proposal	-	+	-	-	-	strengthen
+	+	bend	up	the credit card	-	+	+	-	-	bend completely
+	-	bind	up	the wound	+	+	+	-	-	put bandage on
+	+	block	up	the sink	-	+	+	+	-	obstruct
+	+	blow	up	the balloons	-	+	-	-	-	inflate
+	+	blow	up	the building	+	+	-	+	+	explode
+	+	blow	up	the photo	-	+	-	-	-	enlarge
+	+	blow	up	the scandal	-	+	-	+	-	exaggerate
+	-	boil	up	some water	-	+	+	-	+	boil
+	+	bolster	up	Max	+	+	+	-	-	give hope to
+	+	bolster	up	the theory	-	+	-	-	-	support
+	+	boot	up	the computer	-	+	+	+	+	start

NooJ

Phrasal verb local grammar



Dictionary (.dic file) and Grammar (.nog file) have same name (*phrasal verbs*) and are stored under Lexical Analysis



NooJ can identify phrasal verbs in a variety of texts

Average accuracy of 84.5%

TEXT	Number of examples	Correct continuous examples	Correct discontinuous examples	Prepositions	Misidentifications	Percentage of incorrect	Percentage of correctly identified phrasal verbs
<i>Portrait of a Lady</i>	583	405	83	50	45	16%	84%
Lewis Carroll	294	222	29	24	19	15%	85%
<i>Great Expectations</i>	769	536	108	54	71	16%	84%
Larry King Live (oral) January 2000	584	404	97	51	31	14%	86%
TOTAL	2230	1567	317	179	166	15.5%	84.5%

To improve recall

→ build Lexicon-Grammar tables of purely intransitive phrasal verbs

- MANY intransitive phrasal verbs accept adverb insertion
 - The house **burned** completely **down**
- MOST phrasal prepositional verbs can be used in shortened form, equivalent to intransitive
 - The player **bailed out on** the team
 - The player **bailed out**

Sample Lexicon-Grammar of intransitive phrasal verbs

N_0 :: Nhum	N_0 :: N-hum	Example of N_0	Verb	Particle	N_0 Verb	N_0 Verb Adv Part	Synonym
-	+	the soup	boil	over	+	+	overflow
+	+	Max	boil	over	-	-	become angry / violent
-	+	trouble	boil	up	-	-	start suddenly & be uncontrollable
-	+	explosion	boom	out	-	-	sound extremely loud
+	+	the stock	bottom	out	-	+	stop decreasing before going up
-	+	the elevator	break	down	+	+	stop working
+	-	Mary	break	down	-	-	have an emotional / mental collapse
+	-	Max	break	off	-	-	stop speaking suddenly
+	-	the boy	break	out	-	-	get pimples
-	+	malaria	break	out	-	-	develop suddenly

Lexicon-Grammar of phrasal prepositional verbs

N_0 :: Nhum	N_0 :: N-hum	Example of N_0	Verb	Particle	Preposition	Example of N_1	N_1 :: Nhum	N_1 :: N-hum	N_0 Verb	N_0 Verb Part	N_0 Verb Part N_1	N_0 Verb Prep N_1	N_0 Verb N_1	N_1 Verb Part	Synonym
+	-	the accountant	be	off	in	the estimate	-	+	-	+	-	-	-	+	be inaccurate/inexact
+	-	the police	be	on	to	Max	+	+	-	-	-	-	-	-	be aware of / have facts about
+	-	the student	be	through	with	final exams	+	+	-	+	-	-	-	+	be finished (with)
+	+	Max	be	up	for	the party	-	+	-	+	-	-	-	-	be ready for
-	+	the rain	beat	down	on	the car	+	+	-	+	-	+	-	-	fall on
+	-	Max	beg	off	on	the trip	-	+	-	+	+	-	-	-	ask not to do something
+	+	the rocket	blast	off	for	Mars	-	+	-	+	-	-	-	-	take off / leave for a destination
-	+	the sun	blaze	down	on	Max	+	+	+	+	-	+	-	-	burn intensely
+	+	the sofa	blend	in	with	the decor	-	+	-	+	-	+	-	-	mix well (with)
-	+	the dispute	boil	down	to	a misunderstanding	-	+	-	-	-	-	-	-	be essentially reduced to
+	-	the student	bomb	out	of/in	the exam	-	+	+	+	-	-	+	-	fail

Phrasal prepositional verb columns

- **N_o Verb Part** (MOST)
 - The sun **blazed down**
 - The rocket **blasted off**
 - The accountant **was off**
- **N_o Verb Adv Part** (MANY)
 - The accountant **was way off**
 - The student **was through**
 - The student **was completely through**

- **N₀ Verb Part N₁** (VERY FEW)
 - Max **begged off on** the trip
 - Max **begged off** the trip

- **N₀ Verb Prep N₁** (SOME)
 - The sofa **blended in with** the decor
 - The sofa **blended with** the decor

- **N₀ Verb N₁** (VERY FEW)
 - The student **bombed out in** the exam
 - The student **bombed** the exam

- **N₁ Verb Part** (SOME)
 - The accountant **was off in** the estimate
 - The estimate **was off**

Recently added

- About 200 intransitive phrasal verbs
 - 120 purely intransitive
 - 80 phrasal prepositional verbs that can appear as N_o Verb Part
- Results: improves recall
 - that the human race would **die out**
 - Lord Warburton **loomed up** before her
 - ... she repeated, **glancing up** at the windows

Most new noise created by *come to* and *be + Prep* (12 entries)

- I should like to be **in Rome** with you
- We should be **on good** terms again
- It was an old house **in Albany**
- as if he had been **out of breath**
- Some interesting cases
 - But that he **was better off**; though from what
- Some of these (e.g., ***be out of breath***) can be treated as *be + Prep* frozen expressions

- ***come out, come up*** usually good
 - Henrietta had **come out**, shutting the door ...
 - I guess she'll **come right out**
 - I'll **come up** in half an hour
- But ***come to*** creates new noise
 - Well, that's what it **comes to** in the end (GOOD)
 - Every one that comes **to Europe** ... (NOISE)
 - he spent several days without coming **to see** her (NOISE)
- As with transitive phrasal verbs, most noise created by certain entries

- *time in/out* and *hand in/out/down*
 - They **timed in** the employee [= record arrival]
 - The student **handed in** the work/the exam [= return]
- But in corpora so far, ALL CASES involve *time* as a noun: (40 false positives)
 - I spent a lot of time **in the United States**
 - for the first time **in my life**
 - at the same **time panting out**
- ALL CASES entail *hand* as a noun: (50 false positives)
 - I took her hand **in mine**
 - the two go **hand in hand**
 - or would he easily win **hands down**

Others are less categorical showing both correct and incorrect interpretations:

- to **take on** this issue [= phrasal verb]
- and the initial take **on that** [= noun]

- they're **checking out** to find [= phrasal verb]
- make **the check** out to the Larry King [= noun]
- **make** the check **out to** the Larry King
[= transitive phrasal prepositional verb]

- He **checked in** and registered [= phrasal verb]
- a sudden check **in this tone** [= noun]

- *have on/out* → some are good:
 - In fact they **have** helmets **on**. [=wear]
 - I happened to **have** a cross **on**
 - Mayor Ed Koch **has** a great new book **out** [= publish recently]
 - < Mayor Koch's book **is out** (question of support verbs)
- → Others are misidentified:
 - But actually, we **have** found **out** -- I must plead ...
- → Prepositional phrase:
 - We'll also **have on that plane** some representatives
- → Preposition at end a sentence:
 - There is something everyone **has** an opinion **on**.
[= have an opinion on something]

Some depend on the choice of particle:

- *figure out* is generally good:
- 17 *figure out* in LKL including discontinuous:
 - in trying to **figure it out**
- but *figure in* creates noise:
 - is a pretty outspoken figure **in many areas** in NY
 - when I beheld a solitary figure **in it**. The figure showed itself aware of me,
 - Touchett did not figure **in the list**

- Likewise, *break up / down / out* are generally good:
 - just to **break it up**?
 - I defy it to **break you up**
 - The play **breaks down**
 - The fire would **break out** again and consume her
- But *break in* creates noise:
 - Brian Peterson got a **break in this case**

Future work

- Continue to add intransitive and phrasal prepositional verbs
- Modify grammar to avoid noise
- Still other categories to add: e.g., transitive phrasal prepositional verbs
 - Another new table: N_0 V N_1 Part Prep N_2
 - The teacher **singled** the students **out for** their bad behavior [= choose someone from others to praise or criticize]
 - The teacher **singled out** the students **for** their bad behavior
 - The teacher **singled out** the students
 - The teacher **singled** the students **out**

Transitive Phrasal Prepositional OR Phrasal verb + Prepositional phrase

- Max **turned** the student **on to** NooJ
[= make someone interested in something]
- NooJ **turned** the student **on**
- Max's explanations **turned on** the student **to**
NooJ
- Max's explanations **turned on** the student

Conclusions

- Add exhaustive list of intransitive phrasal verbs to NooJ dictionary to improve NooJ's recall in identifying all phrasal verbs in large corpora
- Modify local grammar for dictionary entries that produce high rate of noise
- Objective: solve a key problem in Natural Language Processing

Thank you for your attention

Merci beaucoup

شكرا جزيلًا