[paths]

train = null

dev = null

vectors = "en\_core\_web\_lg"

init\_tok2vec = null

[system]

gpu\_allocator = null

seed = 0

[nlp]

lang = "en"

pipeline = ["tok2vec","ner"]

batch\_size = 1000

disabled = []

before\_creation = null

after\_creation = null

after\_pipeline\_creation = null

tokenizer = {"@tokenizers":"spacy.Tokenizer.v1"}

vectors = {"@vectors":"spacy.Vectors.v1"}

[components]

[components.ner]

factory = "ner"

incorrect\_spans\_key = null

moves = null

scorer = {"@scorers":"spacy.ner\_scorer.v1"}

update\_with\_oracle\_cut\_size = 100

[components.ner.model]

@architectures = "spacy.TransitionBasedParser.v2"

state\_type = "ner"

extra\_state\_tokens = false

hidden\_width = 64

maxout\_pieces = 2

use\_upper = true

nO = null

[components.ner.model.tok2vec]

@architectures = "spacy.Tok2VecListener.v1"

width = ${components.tok2vec.model.encode.width}

upstream = "\*"

[components.tok2vec]

factory = "tok2vec"

[components.tok2vec.model]

@architectures = "spacy.Tok2Vec.v2"

[components.tok2vec.model.embed]

@architectures = "spacy.MultiHashEmbed.v2"

width = ${components.tok2vec.model.encode.width}

attrs = ["NORM","PREFIX","SUFFIX","SHAPE"]

rows = [5000,1000,2500,2500]

include\_static\_vectors = true

[components.tok2vec.model.encode]

@architectures = "spacy.MaxoutWindowEncoder.v2"

width = 256

depth = 8

window\_size = 1

maxout\_pieces = 3

[corpora]

[corpora.dev]

@readers = "spacy.Corpus.v1"

path = ${paths.dev}

max\_length = 0

gold\_preproc = false

limit = 0

augmenter = null

[corpora.train]

@readers = "spacy.Corpus.v1"

path = ${paths.train}

max\_length = 0

gold\_preproc = false

limit = 0

augmenter = null

[training]

dev\_corpus = "corpora.dev"

train\_corpus = "corpora.train"

seed = ${system.seed}

gpu\_allocator = ${system.gpu\_allocator}

dropout = 0.1

accumulate\_gradient = 1

patience = 1600

max\_epochs = 0

max\_steps = 20000

eval\_frequency = 200

frozen\_components = []

annotating\_components = []

before\_to\_disk = null

before\_update = null

[training.batcher]

@batchers = "spacy.batch\_by\_words.v1"

discard\_oversize = false

tolerance = 0.2

get\_length = null

[training.batcher.size]

@schedules = "compounding.v1"

start = 100

stop = 1000

compound = 1.001

t = 0.0

[training.logger]

@loggers = "spacy.ConsoleLogger.v1"

progress\_bar = false

[training.optimizer]

@optimizers = "Adam.v1"

beta1 = 0.9

beta2 = 0.999

L2\_is\_weight\_decay = true

L2 = 0.01

grad\_clip = 1.0

use\_averages = false

eps = 0.00000001

learn\_rate = 0.001

[training.score\_weights]

ents\_f = 1.0

ents\_p = 0.0

ents\_r = 0.0

ents\_per\_type = null

[pretraining]

[initialize]

vectors = ${paths.vectors}

init\_tok2vec = ${paths.init\_tok2vec}

vocab\_data = null

lookups = null

before\_init = null

after\_init = null

[initialize.components]

[initialize.tokenizer]