

Machine Translation

Natural Language Processing

University of Maryland

Nucleus Sampling

Agenda

The Sentence We Want to Translate

Er sagte mir, 'lass uns blaumachen', aber ich verstand nur Bahnhof

- IBM Model 1
- Phrase-based
- Nucleus Decoding

IBM Model 1 Word Translation Probabilities

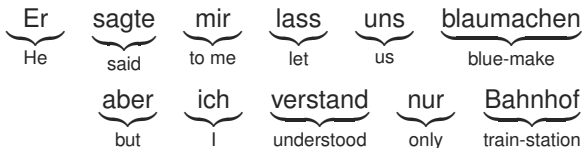
German Word	Top Translation Candidates $p(e g)$
NULL	—
Er	he (0.70), it (0.10), they (0.08), someone (0.06), that (0.06)
sagte	said (0.55), told (0.20), stated (0.10), mentioned (0.08), spoke (0.07)
mir	me (0.75), myself (0.08), with (0.07), to (0.05), personally (0.05)
lass	let (0.60), allow (0.18), leave (0.10), drop (0.07), permit (0.05)
uns	us (0.80), we (0.07), ourselves (0.06), together (0.04), one (0.03)
blaumachen	skip (0.35), slack (0.25), loaf (0.20), ditch (0.10), shirk (0.10)
aber	but (0.70), however (0.15), though (0.08), yet (0.04), nevertheless (0.03)
ich	I (0.90), me (0.05), myself (0.03), one (0.02)
verstand	understood (0.65), got (0.15), grasped (0.10), knew (0.05), followed (0.05)
nur	only (0.70), just (0.18), merely (0.07), simply (0.05)
Bahnhof	station (0.65), train (0.15), terminal (0.08), rail (0.07), platform (0.05)

Phrase-Based Translation: Literal / Small Phrases

German: Er sagte mir, “lass uns blaumachen”, aber ich verstand nur Bahnhof.

English: He said to me, “let us blue-make,” but I understood only train-station.

Word-level Phrase Alignments:








Phrase-Based Translation: Medium / Idiomatic Phrases

German: Er sagte mir, “lass uns blaumachen”, aber ich verstand nur Bahnhof.

English: He told me, “let’s skip work,” but I didn’t understand a thing.

Phrase Alignments:

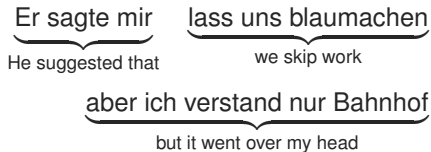
Er sagte mir	lass uns blaumachen	aber ich
		
He told me	let’s skip work	but I
	verstand nur	Bahnhof
		
	only understood	train station

Phrase-Based Translation: Longer / Fluent Phrases

German: Er sagte mir, “lass uns blaumachen”, aber ich verstand nur Bahnhof.

English: He suggested that we skip work, but it all went over my head.

Phrase Alignments:



Step 1 ($p=0.20$)

Prefix:

Token	Prob.	Cum. Prob.
He	0.43	
She	0.33	
Somebody	0.12	
They	0.10	
It	0.01	

Step 1 ($p=0.20$)

Prefix:

Token	Prob.	Cum. Prob.
He	0.43	0.43
She	0.33	0.77
Somebody	0.12	0.89
They	0.10	0.99
It	0.01	1.00

Step 1 ($p=0.20$)

Prefix:

Token	Prob.	Cum. Prob.
He	0.43	0.43
She	0.33	0.77
Somebody	0.12	0.89
They	0.10	0.99
It	0.01	1.00

{He}

Step 1 ($p=0.20$)

Prefix:

Token	Prob.	Cum. Prob.
He	0.43	0.43
She	0.33	0.77
Somebody	0.12	0.89
They	0.10	0.99
It	0.01	1.00

{He}

Step 2 ($p=0.20$)

Prefix: He

Token	Prob.	Cum. Prob.
said	0.55	
asked	0.16	
told	0.15	
mentioned	0.09	
explained	0.05	
spoke	0.01	

Step 2 ($p=0.20$)

Prefix: He

Token	Prob.	Cum. Prob.
said	0.55	0.55
asked	0.16	0.71
told	0.15	0.86
mentioned	0.09	0.95
explained	0.05	0.99
spoke	0.01	1.00

Step 2 ($p=0.20$)

Prefix: He

Token	Prob.	Cum. Prob.
said	0.55	0.55
asked	0.16	0.71
told	0.15	0.86
mentioned	0.09	0.95
explained	0.05	0.99
spoke	0.01	1.00

{said}

Step 2 ($p=0.20$)

Prefix: He

Token	Prob.	Cum. Prob.
said	0.55	0.55
asked	0.16	0.71
told	0.15	0.86
mentioned	0.09	0.95
explained	0.05	0.99
spoke	0.01	1.00

{said}

Step 3 ($p=0.20$)

Prefix: He said

Token	Prob.	Cum. Prob.
me	0.55	
nobody	0.16	
them	0.14	
him	0.11	
her	0.04	

Step 3 ($p=0.20$)

Prefix: He said

Token	Prob.	Cum. Prob.
me	0.55	0.55
nobody	0.16	0.71
them	0.14	0.84
him	0.11	0.96
her	0.04	1.00

Step 3 ($p=0.20$)

Prefix: He said

Token	Prob.	Cum. Prob.
me	0.55	0.55
nobody	0.16	0.71
them	0.14	0.84
him	0.11	0.96
her	0.04	1.00

{me}

Step 3 ($p=0.20$)

Prefix: He said

Token	Prob.	Cum. Prob.
me	0.55	0.55
nobody	0.16	0.71
them	0.14	0.84
him	0.11	0.96
her	0.04	1.00

{me}

Step 4 ($p=0.20$)

Prefix: He said me

Token	Prob.	Cum. Prob.
maybe	0.27	
why	0.20	
'we	0.14	
let	0.14	
'should	0.11	
let's	0.09	
how	0.05	

Step 4 ($p=0.20$)

Prefix: He said me

Token	Prob.	Cum. Prob.
maybe	0.27	0.27
why	0.20	0.47
'we	0.14	0.61
let	0.14	0.75
'should	0.11	0.86
let's	0.09	0.95
how	0.05	1.00

Step 4 ($p=0.20$)

Prefix: He said me

Token	Prob.	Cum. Prob.
maybe	0.27	0.27
why	0.20	0.47
'we	0.14	0.61
let	0.14	0.75
'should	0.11	0.86
let's	0.09	0.95
how	0.05	1.00

{let, let's}

Step 4 ($p=0.20$)

Prefix: He said me

Token	Prob.	Cum. Prob.
maybe	0.27	0.27
why	0.20	0.47
'we	0.14	0.61
let	0.14	0.75
'should	0.11	0.86
let's	0.09	0.95
how	0.05	1.00

{let, let's}

Step 5 ($p=0.20$)

Prefix: He said me let

Token	Prob.	Cum. Prob.
us	0.40	
them	0.23	
you	0.21	
nobody	0.10	
everyone	0.06	

Step 5 ($p=0.20$)

Prefix: He said me let

Token	Prob.	Cum. Prob.
us	0.40	0.40
them	0.23	0.62
you	0.21	0.83
nobody	0.10	0.94
everyone	0.06	1.00

Step 5 ($p=0.20$)

Prefix: He said me let

Token	Prob.	Cum. Prob.
us	0.40	0.40
them	0.23	0.62
you	0.21	0.83
nobody	0.10	0.94
everyone	0.06	1.00

{us}

Step 5 ($p=0.20$)

Prefix: He said me let

Token	Prob.	Cum. Prob.
us	0.40	0.40
them	0.23	0.62
you	0.21	0.83
nobody	0.10	0.94
everyone	0.06	1.00

{us}

Step 6 ($p=0.20$)

Prefix: He said me let us

Token	Prob.	Cum. Prob.
skip	0.39	
cut	0.23	
ditch	0.13	
miss	0.11	
escape	0.08	
avoid	0.06	

Step 6 ($p=0.20$)

Prefix: He said me let us

Token	Prob.	Cum. Prob.
skip	0.39	0.39
cut	0.23	0.62
ditch	0.13	0.75
miss	0.11	0.86
escape	0.08	0.94
avoid	0.06	1.00

Step 6 ($p=0.20$)

Prefix: He said me let us

Token	Prob.	Cum. Prob.
skip	0.39	0.39
cut	0.23	0.62
ditch	0.13	0.75
miss	0.11	0.86
escape	0.08	0.94
avoid	0.06	1.00

{skip, cut}

Step 6 ($p=0.20$)

Prefix: He said me let us

Token	Prob.	Cum. Prob.
skip	0.39	0.39
cut	0.23	0.62
ditch	0.13	0.75
miss	0.11	0.86
escape	0.08	0.94
avoid	0.06	1.00

{skip, cut}

Step 7 ($p=0.20$)

Prefix: He said me let us cut

Token	Prob.	Cum. Prob.
work	0.88	
class	0.04	
today	0.03	
school	0.03	
it	0.03	

Step 7 ($p=0.20$)

Prefix: He said me let us cut

Token	Prob.	Cum. Prob.
work	0.88	0.88
class	0.04	0.91
today	0.03	0.95
school	0.03	0.97
it	0.03	1.00

Step 7 ($p=0.20$)

Prefix: He said me let us cut

Token	Prob.	Cum. Prob.
work	0.88	0.88
class	0.04	0.91
today	0.03	0.95
school	0.03	0.97
it	0.03	1.00

{work}

Step 7 ($p=0.20$)

Prefix: He said me let us cut

Token	Prob.	Cum. Prob.
work	0.88	0.88
class	0.04	0.91
today	0.03	0.95
school	0.03	0.97
it	0.03	1.00

{work}

Step 8 ($p=0.20$)

Prefix: He said me let us cut work

Token	Prob.	Cum. Prob.
but	0.46	
so	0.22	
because	0.15	
though	0.10	
and	0.07	

Step 8 ($p=0.20$)

Prefix: He said me let us cut work

Token	Prob.	Cum. Prob.
but	0.46	0.46
so	0.22	0.68
because	0.15	0.83
though	0.10	0.93
and	0.07	1.00

Step 8 ($p=0.20$)

Prefix: He said me let us cut work

Token	Prob.	Cum. Prob.
but	0.46	0.46
so	0.22	0.68
because	0.15	0.83
though	0.10	0.93
and	0.07	1.00

{but}

Step 8 ($p=0.20$)

Prefix: He said me let us cut work

Token	Prob.	Cum. Prob.
but	0.46	0.46
so	0.22	0.68
because	0.15	0.83
though	0.10	0.93
and	0.07	1.00

{but}

Step 9 ($p=0.20$)

Prefix: He said me let us cut work but

Token	Prob.	Cum. Prob.
I	0.37	
nobody	0.23	
someone	0.18	
they	0.12	
we	0.10	

Step 9 ($p=0.20$)

Prefix: He said me let us cut work but

Token	Prob.	Cum. Prob.
I	0.37	0.37
nobody	0.23	0.59
someone	0.18	0.78
they	0.12	0.90
we	0.10	1.00

Step 9 ($p=0.20$)

Prefix: He said me let us cut work but

Token	Prob.	Cum. Prob.
I	0.37	0.37
nobody	0.23	0.59
someone	0.18	0.78
they	0.12	0.90
we	0.10	1.00

{I}

Step 9 ($p=0.20$)

Prefix: He said me let us cut work but

Token	Prob.	Cum. Prob.
I	0.37	0.37
nobody	0.23	0.59
someone	0.18	0.78
they	0.12	0.90
we	0.10	1.00

{**I**}

Step 10 ($p=0.20$)

Prefix: He said me let us cut work but I

Token	Prob.	Cum. Prob.
-------	-------	------------

only	0.36	
------	------	--

never	0.23	
-------	------	--

just	0.16	
------	------	--

barely	0.12	
--------	------	--

didn't	0.11	
--------	------	--

hardly	0.02	
--------	------	--

Step 10 ($p=0.20$)

Prefix: He said me let us cut work but I

Token	Prob.	Cum. Prob.
only	0.36	0.36
never	0.23	0.59
just	0.16	0.75
barely	0.12	0.87
didn't	0.11	0.98
hardly	0.02	1.00

Step 10 ($p=0.20$)

Prefix: He said me let us cut work but I

Token	Prob.	Cum. Prob.
only	0.36	0.36
never	0.23	0.59
just	0.16	0.75
barely	0.12	0.87
didn't	0.11	0.98
hardly	0.02	1.00

{only, didn't}

Step 10 ($p=0.20$)

Prefix: He said me let us cut work but I

Token	Prob.	Cum. Prob.
only	0.36	0.36
never	0.23	0.59
just	0.16	0.75
barely	0.12	0.87
didn't	0.11	0.98
hardly	0.02	1.00

{only, didn't}

Step 11 ($p=0.20$)

Prefix: He said me let us cut work but I
didn't

Token	Prob.	Cum. Prob.
understand	0.36	
got	0.27	
followed	0.15	
heard	0.12	
caught	0.07	
knew	0.02	

Step 11 ($p=0.20$)

Prefix: He said me let us cut work but I
didn't

Token	Prob.	Cum. Prob.
understand	0.36	0.36
got	0.27	0.63
followed	0.15	0.79
heard	0.12	0.91
caught	0.07	0.98
knew	0.02	1.00

Step 11 ($p=0.20$)

Prefix: He said me let us cut work but I
didn't

Token	Prob.	Cum. Prob.
understand	0.36	0.36
got	0.27	0.63
followed	0.15	0.79
heard	0.12	0.91
caught	0.07	0.98
knew	0.02	1.00

{understand, got}

Step 11 ($p=0.20$)

Prefix: He said me let us cut work but I
didn't

Token	Prob.	Cum. Prob.
understand	0.36	0.36
got	0.27	0.63
followed	0.15	0.79
heard	0.12	0.91
caught	0.07	0.98
knew	0.02	1.00

{understand, got}

Step 12 ($p=0.20$)

Prefix: He said me let us cut work but I
didn't got

Token	Prob.	Cum. Prob.
nothing	0.57	
zero	0.16	
zilch	0.12	
anything	0.09	
something	0.06	

Step 12 ($p=0.20$)

Prefix: He said me let us cut work but I
didn't got

Token	Prob.	Cum. Prob.
nothing	0.57	0.57
zero	0.16	0.73
zilch	0.12	0.85
anything	0.09	0.94
something	0.06	1.00

Step 12 ($p=0.20$)

Prefix: He said me let us cut work but I
didn't got

Token	Prob.	Cum. Prob.
nothing	0.57	0.57
zero	0.16	0.73
zilch	0.12	0.85
anything	0.09	0.94
something	0.06	1.00

{nothing}

Step 12 ($p=0.20$)

Prefix: He said me let us cut work but I
didn't got

Token	Prob.	Cum. Prob.
nothing	0.57	0.57
zero	0.16	0.73
zilch	0.12	0.85
anything	0.09	0.94
something	0.06	1.00

{nothing}

Step 1 ($p=0.50$)

Prefix:

Token	Prob.	Cum. Prob.
It	0.32	
He	0.28	
Somebody	0.28	
This	0.06	
The	0.06	
They	0.01	

Step 1 ($p=0.50$)

Prefix:

Token	Prob.	Cum. Prob.
It	0.32	0.32
He	0.28	0.60
Somebody	0.28	0.88
This	0.06	0.93
The	0.06	0.99
They	0.01	1.00

Step 1 ($p=0.50$)

Prefix:

Token	Prob.	Cum. Prob.
It	0.32	0.32
He	0.28	0.60
Somebody	0.28	0.88
This	0.06	0.93
The	0.06	0.99
They	0.01	1.00

{He, It, Somebody}

Step 1 ($p=0.50$)

Prefix:

Token	Prob.	Cum. Prob.
It	0.32	0.32
He	0.28	0.60
Somebody	0.28	0.88
This	0.06	0.93
The	0.06	0.99
They	0.01	1.00

{He, It, Somebody}

Step 2 ($p=0.50$)

Prefix: He

Token	Prob.	Cum. Prob.
proposed	0.27	
asked	0.18	
said	0.17	
remarked	0.16	
told	0.13	
suggested	0.08	

Step 2 ($p=0.50$)

Prefix: He

Token	Prob.	Cum. Prob.
proposed	0.27	0.27
asked	0.18	0.45
said	0.17	0.63
remarked	0.16	0.78
told	0.13	0.92
suggested	0.08	1.00

Step 2 ($p=0.50$)

Prefix: He

Token	Prob.	Cum. Prob.
proposed	0.27	0.27
asked	0.18	0.45
said	0.17	0.63
remarked	0.16	0.78
told	0.13	0.92
suggested	0.08	1.00

{asked, said, proposed}

Step 2 ($p=0.50$)

Prefix: He

Token	Prob.	Cum. Prob.
proposed	0.27	0.27
asked	0.18	0.45
said	0.17	0.63
remarked	0.16	0.78
told	0.13	0.92
suggested	0.08	1.00

{asked, said, proposed}

Step 3 ($p=0.50$)

Prefix: He asked

Token	Prob.	Cum. Prob.
me	0.38	
how	0.29	
to	0.17	
something	0.07	
that	0.05	
about	0.02	
if	0.02	

Step 3 ($p=0.50$)

Prefix: He asked

Token	Prob.	Cum. Prob.
me	0.38	0.38
how	0.29	0.67
to	0.17	0.84
something	0.07	0.91
that	0.05	0.96
about	0.02	0.98
if	0.02	1.00

Step 3 ($p=0.50$)

Prefix: He asked

Token	Prob.	Cum. Prob.
me	0.38	0.38
how	0.29	0.67
to	0.17	0.84
something	0.07	0.91
that	0.05	0.96
about	0.02	0.98
if	0.02	1.00

{to, me, how}

Step 3 ($p=0.50$)

Prefix: He asked

Token	Prob.	Cum. Prob.
me	0.38	0.38
how	0.29	0.67
to	0.17	0.84
something	0.07	0.91
that	0.05	0.96
about	0.02	0.98
if	0.02	1.00

{to, me, how}

Step 4 ($p=0.50$)

Prefix: He asked how

Token	Prob.	Cum. Prob.
would	0.34	
sorta	0.31	
'bout	0.19	
perhaps	0.08	
kinda	0.08	
maybe	0.00	

Step 4 ($p=0.50$)

Prefix: He asked how

Token	Prob.	Cum. Prob.
would	0.34	0.34
sorta	0.31	0.65
'bout	0.19	0.84
perhaps	0.08	0.92
kinda	0.08	1.00
maybe	0.00	1.00

Step 4 ($p=0.50$)

Prefix: He asked how

Token	Prob.	Cum. Prob.
would	0.34	0.34
sorta	0.31	0.65
'bout	0.19	0.84
perhaps	0.08	0.92
kinda	0.08	1.00
maybe	0.00	1.00

{'bout, would, maybe}

Step 4 ($p=0.50$)

Prefix: He asked how

Token	Prob.	Cum. Prob.
would	0.34	0.34
sorta	0.31	0.65
'bout	0.19	0.84
perhaps	0.08	0.92
kinda	0.08	1.00
maybe	0.00	1.00

{'bout, would, maybe}

Step 5 ($p=0.50$)

Prefix: He asked how 'bout

Token	Prob.	Cum. Prob.
playing	0.21	
shirking	0.20	
skipping	0.18	
cutting	0.17	
avoiding	0.10	
ditching	0.08	
bailing	0.05	

Step 5 ($p=0.50$)

Prefix: He asked how 'bout

Token	Prob.	Cum. Prob.
playing	0.21	0.21
shirking	0.20	0.42
skipping	0.18	0.60
cutting	0.17	0.78
avoiding	0.10	0.87
ditching	0.08	0.95
bailing	0.05	1.00

Step 5 ($p=0.50$)

Prefix: He asked how 'bout

Token	Prob.	Cum. Prob.
playing	0.21	0.21
shirking	0.20	0.42
skipping	0.18	0.60
cutting	0.17	0.78
avoiding	0.10	0.87
ditching	0.08	0.95
bailing	0.05	1.00

{playing, skipping, shirking, avoiding}

Step 5 ($p=0.50$)

Prefix: He asked how 'bout

Token	Prob.	Cum. Prob.
playing	0.21	0.21
shirking	0.20	0.42
skipping	0.18	0.60
cutting	0.17	0.78
avoiding	0.10	0.87
ditching	0.08	0.95
bailing	0.05	1.00

{playing, skipping, shirking, avoiding}

Step 6 ($p=0.50$)

Prefix: He asked how 'bout avoiding

Token	Prob.	Cum. Prob.
class	0.32	
seminar	0.29	
work	0.20	
school	0.11	
responsibility	0.05	
duty	0.04	

Step 6 ($p=0.50$)

Prefix: He asked how 'bout avoiding

Token	Prob.	Cum. Prob.
class	0.32	0.32
seminar	0.29	0.61
work	0.20	0.81
school	0.11	0.91
responsibility	0.05	0.96
duty	0.04	1.00

Step 6 ($p=0.50$)

Prefix: He asked how 'bout avoiding

Token	Prob.	Cum. Prob.
class	0.32	0.32
seminar	0.29	0.61
work	0.20	0.81
school	0.11	0.91
responsibility	0.05	0.96
duty	0.04	1.00

{work, class, seminar}

Step 6 ($p=0.50$)

Prefix: He asked how 'bout avoiding

Token	Prob.	Cum. Prob.
class	0.32	0.32
seminar	0.29	0.61
work	0.20	0.81
school	0.11	0.91
responsibility	0.05	0.96
duty	0.04	1.00

{work, class, seminar}

Step 7 ($p=0.50$)

Prefix: He asked how 'bout avoiding class

Token	Prob.	Cum. Prob.
however	0.34	
yet	0.33	
though	0.13	
but	0.11	
except	0.08	
still	0.01	

Step 7 ($p=0.50$)

Prefix: He asked how 'bout avoiding class

Token	Prob.	Cum. Prob.
however	0.34	0.34
yet	0.33	0.67
though	0.13	0.80
but	0.11	0.91
except	0.08	0.99
still	0.01	1.00

Step 7 ($p=0.50$)

Prefix: He asked how 'bout avoiding class

Token	Prob.	Cum. Prob.
however	0.34	0.34
yet	0.33	0.67
though	0.13	0.80
but	0.11	0.91
except	0.08	0.99
still	0.01	1.00

{but, however, though}

Step 7 ($p=0.50$)

Prefix: He asked how 'bout avoiding class

Token	Prob.	Cum. Prob.
however	0.34	0.34
yet	0.33	0.67
though	0.13	0.80
but	0.11	0.91
except	0.08	0.99
still	0.01	1.00

{**but**, however, though}

Step 8 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but

Token	Prob.	Cum. Prob.
the	0.33	
it	0.23	
that	0.12	
seemed	0.11	
he	0.10	
I	0.08	
everything	0.02	

Step 8 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but

Token	Prob.	Cum. Prob.
the	0.33	0.33
it	0.23	0.57
that	0.12	0.68
seemed	0.11	0.80
he	0.10	0.89
I	0.08	0.98
everything	0.02	1.00

Step 8 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but

Token	Prob.	Cum. Prob.
the	0.33	0.33
it	0.23	0.57
that	0.12	0.68
seemed	0.11	0.80
he	0.10	0.89
I	0.08	0.98
everything	0.02	1.00

{it, the, seemed}

Step 8 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but

Token	Prob.	Cum. Prob.
the	0.33	0.33
it	0.23	0.57
that	0.12	0.68
seemed	0.11	0.80
he	0.10	0.89
I	0.08	0.98
everything	0.02	1.00

{it, the, seemed}

Step 9 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it

Token	Prob.	Cum. Prob.
seemed	0.62	
was	0.26	
felt	0.10	
appeared	0.01	
sounded	0.01	

Step 9 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it

Token	Prob.	Cum. Prob.
seemed	0.62	0.62
was	0.26	0.89
felt	0.10	0.98
appeared	0.01	0.99
sounded	0.01	1.00

Step 9 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it

Token	Prob.	Cum. Prob.
seemed	0.62	0.62
was	0.26	0.89
felt	0.10	0.98
appeared	0.01	0.99
sounded	0.01	1.00

{was, seemed}

Step 9 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it

Token	Prob.	Cum. Prob.
seemed	0.62	0.62
was	0.26	0.89
felt	0.10	0.98
appeared	0.01	0.99
sounded	0.01	1.00

{was, seemed}

Step 10 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed

Token	Prob.	Cum. Prob.
unintelligible	0.26	
Greek	0.21	
unclear	0.21	
weird	0.15	
confusing	0.13	
nonsense	0.03	
lost	0.01	

Step 10 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed

Token	Prob.	Cum. Prob.
unintelligible	0.26	0.26
Greek	0.21	0.47
unclear	0.21	0.68
weird	0.15	0.83
confusing	0.13	0.96
nonsense	0.03	0.99
lost	0.01	1.00

Step 10 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed

Token	Prob.	Cum. Prob.
unintelligible	0.26	0.26
Greek	0.21	0.47
unclear	0.21	0.68
weird	0.15	0.83
confusing	0.13	0.96
nonsense	0.03	0.99
lost	0.01	1.00

{Greek, unintelligible, lost, unclear}

Step 10 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed

Token	Prob.	Cum. Prob.
unintelligible	0.26	0.26
Greek	0.21	0.47
unclear	0.21	0.68
weird	0.15	0.83
confusing	0.13	0.96
nonsense	0.03	0.99
lost	0.01	1.00

{Greek, unintelligible, lost, unclear}

Step 11 (p=0.50)

Prefix: He asked how 'bout avoiding class
but it seemed Greek

Token	Prob.	Cum. Prob.
for	0.51	
to	0.26	
around	0.14	
in front of	0.08	
with	0.00	

Step 11 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed Greek

Token	Prob.	Cum. Prob.
for	0.51	0.51
to	0.26	0.77
around	0.14	0.91
in front of	0.08	1.00
with	0.00	1.00

Step 11 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed Greek

Token	Prob.	Cum. Prob.
for	0.51	0.51
to	0.26	0.77
around	0.14	0.91
in front of	0.08	1.00
with	0.00	1.00

{to, for}

Step 11 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed Greek

Token	Prob.	Cum. Prob.
for	0.51	0.51
to	0.26	0.77
around	0.14	0.91
in front of	0.08	1.00
with	0.00	1.00

{to, for}

Step 12 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed Greek for

Token	Prob.	Cum. Prob.
me	0.60	
her	0.12	
everyone	0.12	
them	0.10	
him	0.06	

Step 12 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed Greek for

Token	Prob.	Cum. Prob.
me	0.60	0.60
her	0.12	0.72
everyone	0.12	0.84
them	0.10	0.94
him	0.06	1.00

Step 12 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed Greek for

Token	Prob.	Cum. Prob.
me	0.60	0.60
her	0.12	0.72
everyone	0.12	0.84
them	0.10	0.94
him	0.06	1.00

{me}

Step 12 ($p=0.50$)

Prefix: He asked how 'bout avoiding class
but it seemed Greek for

Token	Prob.	Cum. Prob.
me	0.60	0.60
her	0.12	0.72
everyone	0.12	0.84
them	0.10	0.94
him	0.06	1.00

{me}