Cockney Rhyming Slang Generator

The project seems interesting as can be considered a sample of Natural-language generation and text transformation aspects of NLP.

Rhyming slang is a form of phrase construction in the English language and is especially prevalent in dialectal English from the East End of London; hence the alternative name, Cockney rhyming slang. **The construction involves replacing a common word with a rhyming phrase of two or three words and then, in almost all cases, omitting the secondary rhyming word**, making the origin and meaning of the phrase elusive to listeners not in the know. So, this is a kind of phono-semantic rhyming slang, which includes a semantic link between the slang expression and its referent (the thing it refers to).

There are some realizations of the concept already. One of them presumes that Rhyming slang is generated through a combination of 2 linguistic corpora: the CMU Pronouncing Dictionary and the Edinburgh Associative Corpus. CMU is used to index words according to their rhyming pattern, and EAT is used to group pronounceable words by semantic association.

Once everything is indexed, the actual generation of slang is pretty straightforward. First, take a word and find its rhyme pattern. Then get all of the other words that share this pattern. Randomly go through each rhyming word until one is found that has associated words that meet a particular threshold of relevance. Finally, take one of those words and return it along with the rhyme and the original word.

However, the idea of choosing a certain phono-semantic pattern is of great interest and gives plenty of space for exploration. It requires research to get understanding of such generation performed in natural language. As talking about Cockney rhyming slang one should definitely underline its ironical, jocular and controversial character, that lies in the basis of this generation. Program recreating of such regularity will be the most fascinating part of the task.

In a simple solution the program computationally generates unique rhyming slang for common words, thus creating slang vocabulary or in more complex one also finds and replaces certain words in a given text with a generated slang.

Samples:

Slang vocabulary

Input	Rhyming pattern	Output
Cockney Rhyming Slang	Chitty Chitty Bang Bang	Chitty Chitty
Alone	Todd Sloan	Todd
Braces	Airs and Graces	Airs
Pub	Nuclear Sub	Nuclear
Rain	Pleasure and Pain	Pleasure
Stout	Salmon and Trout	Salmon
Girl	Twist and Twirl	Twist
Dyke (lesbian)	Raleigh Bike	Raleigh
Pakistani	Reg Varney	Reg
Gin	Vera Lynn	Vera
	Needle and Pin	Needle
Tonic	Philharmonic	Phil
Poof (homosexual)	Tin Roof	Tin Roof
Rum	Tom Thumb	Tom
Smoke (cigarette)	Laugh and Joke	Laugh
Jugs (breasts)	Carpets and Rugs	Carpets
Whore	Roger Moore	Roger
Powder (cocaine)	Nikki Lauder	Nikki
Rave (dance)	Comedy Dave	Comedy

Text sample

Input	Output
Let's talk some Cockney Rhyming Slang, little kitty.	Let's talk some chitty chitty, little kitty.
Looks like I was all alone that night.	Looks like I was on my Todd that night.
So I've put new braces on and was off to pub.	So I've put new airs on and was off to nuclear.
It was raining enough for swimming.	Pleasure enough for swimming.
I'm in, got some stout.	I'm in, got some salmon.
She looks like a nice girl, but a right dyke.	She looks like a nice twist, but a right Raleigh.
Oh, Martin's new bird's a Pakistani.	Oh, Martin's new bird's a Reg.
Got some gin.	Got some Vera and Phil.
I think he might be a poof.	I think he might be a tin roof.
Got a drop of rum.	Got a drop of Tom.
Going out for a quick smoke.	Going out for a quick laugh.
See a nice set of jugs.	See a nice set of carpets.
Should buy her some gin.	Should buy her some needle.
I was trying to get my trousers back on, and the dirty whore is running up the street with my wallet.	I was trying to get my trousers back on, and the dirty roger is running up the street with my wallet.
Finally, need to get some powder and go to rave.	Finally, need to get some Nikki and go to comedy.