Grokking Delay in epoch 10.6 298K -11.1 95K -11.1 12.4 0.1

a)					
Siz	2614		22.0	102.4	045.5
set	26K -		33.0	192.4	845.5
ţ					
æ					

Irain Dataset	26K -				33.0	192.4	845.5
	12K -			27.6	214.1	945.6	2.0k
=	2.8K -	12.6	47.6	364.9	855.0	3.7k	10.2k

Irain	12K -			27.6	214.1	945.6	2.UK
ļ	2.8K -	12.6	47.6	364.9	855.0	3.7k	10.2k
	0.614	250.6	400.6	2.41	7.01		22.01

32

Batch Size

128

512

1024

8